



Puiden kuolleisuuden seuraaminen monivuotisten kaukokartoitusaineistojen avulla

Aleksi Kinnunen
Maisterintutkielma
Helsingin yliopisto
Metsätieteiden maisteriohjelma
Metsien ekologia ja käyttö
Huhtikuu 2021

Tiedekunta/Osasto Fakultet/Sektion – Faculty Maatalous-metsätieteellinen tiedekunta		Laitos/Institution– Department Metsätieteiden osasto, Metsätieteiden maisteriohjelma
Tekijä/Författare – Author Aleksi Kinnunen		
Työn nimi / Arbetets titel – Title Puiden kuolleisuuden seuraaminen monivuotisten kaukokartoitusaineistojen avulla		
Oppiaine /Läroämne – Subject Metsien ekologia ja käyttö		
Työn laji/Arbetets art – Level Maisterintutkielma	Aika/Datum – Month and year Huhtikuu 2021	Sivumäärä/ Sidoantal – Number of pages 65
<p>Tiivistelmä/Referat – Abstract</p> <p>Puut kohtaavat yhä moninaisempia terveysuhkia. Ilmastonmuutoksen kokonaisvaikutuksia puihin ja metsiin on vaikea ennustaa. Lämpenevän ilmaston seurauksena kasvukausi pitenee parantaen puuston kasvua, mutta samalla kuivuus- ja hyönteistuhot saatavat yleistyä ja myrskytuhojen riski kasvaa. Puiden kuolleisuuden seurannalla on monia hyötyjä. Sillä voidaan arvioida metsien terveydentilaa, tuottavuutta, hiilensidontaa sekä kuolleiden puiden ekologisia vaikutuksia metsäekosysteemeissä. Puun kuolemaan johtavia syitä voivat olla muun muassa biologiset, ilmastolliset tai ihmiseen liittyvät tekijät. Seuranta voi lisätä ymmärrystä kuolemaan johtavista syistä ja mahdollisesti auttaa metsien paremmassa suojelussa. Puihin liittyvä kuolleisuus on spatiaalisesti ja ajallisesti epäsäännöllinen prosessi, jonka seuraaminen on hankalaa perinteisin inventointimenetelmin. Kaukokartoitus mahdollistaakin puiden kuolleisuuden tehokkaamman kartoittamisen ja seurannan.</p> <p>Tämän tutkielman tarkoituksena oli selvittää, miten kaukokartoitusaineistoja voidaan hyödyntää puiden kuolleisuuden seurannassa. Tavoitteena oli saada selville, miten puiden kuolleisuus on vaihdellut alueellisesti ja määrällisesti Helsingin keskuspuistossa sekä millä tarkkuudella kuolleita puita voidaan tunnistaa ilmakuvilta. Tutkimusjaksoksi valittiin 2005–2019, jonka ajalta tarkasteltiin seitsemän eri vuoden ortoilmakuvia. 15 vuoden ajanjakson kattavalta ilmakuva-aikasarjalta kerättiin visuaalisella kuvatulkinnalla 14 212 puun referenssiaineisto. Aineisto sisälsi terveitä, sairaita sekä kuolleita puita. Sairaiden ja kuolleiden puiden sijainneista luotiin lämpökartta-aikasarjat, joilla tarkasteltiin kuolleisuuden määrällistä ja alueellista vaihtelua. Vuosien keskilämpötiloja sekä sademääriä verrattiin kuolleiden puiden määriin ja tarkasteltiin havaintojen välisiä korrelaatioita. Kerättyä referenssiaineistoa hyödynnettiin myös terveydentilaluokittelussa, jotka toteutettiin puoliatomaattisilla koneoppimismenetelmillä. Luokittelussa pyrittiin tunnistamaan terveet, sairaat sekä kuolleet puut mahdollisimman hyvin toisistaan. Yksittäisten puiden latvukset rajattiin laserkeilausaineistosta saaduilla latvussegmenteillä. Rajattujen latvusten sisältämistä pikseleistä laskettiin yksittäisille puille niitä kuvaavia kuvapiirteitä.</p> <p>Puiden kuolleisuudessa havaittiin tapahtuneen huomattavia muutoksia. Tutkimusjakson alussa kuolleiden puiden määrät kasvoivat vuodesta 2005 vuoteen 2009 merkittävästi. Mahdolliseksi syyksi tunnistettiin vuoden 2006 poikkeuksellisen kuiva kesä. Seuraavina vuosina tilanne pysyi maltillisena, mutta määrällisesti sekä alueellisesti laajimmillaan kuolleisuus oli vuonna 2017. Kokonaisuutena kuolleisuudessa oli havaittavissa nouseva trendi tutkimusjakson aikana, ja vuosien keskilämpötilojen havaittiin korreloivan vahvasti kuolleiden puiden määrän kanssa ($r=0,73$). Puiden terveydentilan luokittelutarkkuudet vaihtelivat vuosittain 89–96 % välillä. Seitsemän vuoden keskiarvotarkkuus oli 93,6 % kappa-arvon ollessa 0,88. Luokittelussa tärkeimmiksi piirteiksi osoittautuivat sinisestä kanavasta lasketut piirteet, kuten kanavan maksimiarvo (B_{max}), kanavan maksimin ja minimin erotus (B_{range}) sekä jakauman vinous (B_{skew}).</p> <p>Tutkielman tulokset osoittivat, että puiden kuolleisuutta voidaan seurata kaukokartoitusaineistoilla. Kuolleiden puiden määrissä havaittiin selkeitä muutoksia aikasarjan avulla ja mahdollisia syitä onnistuttiin tunnistamaan. Tunnistamalla syitä kasvavan kuolleisuuden taustalla voidaan myös ymmärtää paremmin ilmastomuutoksen vaikutuksia. Puiden terveydentilan luokittelutarkkuudet olivat hyvällä tasolla ja kuolleita puita voidaan kartoittaa ilmakuvilta puoliatomaattisilla menetelmillä. Tutkielman perusteella voidaan perustellusti todeta, että ilmakuva-aikasarjoilla voidaan havainnoida muutoksia puiden kuolleisuudessa. Lisäksi voidaan todeta kuolleiden puiden puoliatomaattisen tunnistamisen ilmakuvilta olevan riittävän tarkkaa laajamittaiseen käyttöön.</p>		
<p>Avainsanat – Nyckelord – Keywords Puiden kuolleisuus, kaukokartoitus, ilmastomuutos, puiden terveys, koneoppiminen</p>		
<p>Säilytyspaikka – Förvaringställe – Where deposited Helsingin yliopiston kirjasto – Helda / E-thesis (opinnäytteet) ethesis.helsinki.fi</p>		
<p>Muita tietoja – Övriga uppgifter – Additional information</p>		

Tiedekunta/Osasto Fakultet/Sektion – Faculty Faculty of Agriculture and Forestry		Laitos/Institution– Department The Department of Forest Sciences, Master's Programme in Forest Sciences	
Tekijä/Författare – Author Aleksi Kinnunen			
Työn nimi / Arbetets titel – Title Monitoring tree mortality using multiyear remote sensing data			
Oppiaine /Läroämne – Subject Forest Ecology and Management			
Työn laji/Arbetets art – Level Master's thesis	Aika/Datum – Month and year April 2021	Sivumäärä/ Sidoantal – Number of pages 65	
<p>Tiivistelmä/Referat – Abstract</p> <p>Trees face an increasing variety of health threats. The overall effects of climate change on trees and forests are difficult to predict. As a result of the warming climate, the growing season is lengthening, improving the growth of the trees, but at the same time drought and insect damages may become more common and the risk of storm damage increases. There are many benefits to monitoring tree mortality. It can be used to assess the health status of forests, productivity, carbon sequestration and the ecological impacts of dead trees on forest ecosystems. Causes leading to tree death can include biological, climatic or human related factors. Monitoring can increase understanding of the causes of death and potentially help to protect forests better. Tree-related mortality is a spatially and temporally irregular process that is difficult to monitor using traditional inventory methods. Remote sensing makes it possible to map and monitor tree mortality more effectively.</p> <p>The purpose of this thesis was to find out how remote sensing data can be utilized in monitoring tree mortality. The aim was to find out how tree mortality has varied regionally and quantitatively in the Central Park of Helsinki and how accurately dead trees can be identified from aerial imagery. The study period was 2005–2019, during which orthophotos of seven different years were examined. Reference data of 14 212 trees were collected from the aerial time series covering a 15-year period by visual image interpretation. The data included healthy, weakened and dead trees. Heatmap time series were created from the locations of weakened and dead trees to examine quantitative and regional variability in mortality. The average temperatures over the years as well as the rainfall were compared with the dead tree numbers and the correlations between the observations were examined. The collected reference data was also utilized in health status classifications, which were implemented using semi-automatic machine learning methods. The object of the classifications was to identify healthy, weakened and dead trees as well as possible from each other. The canopies of individual trees were delimited by canopy segments obtained from laser scanning data. From the pixels contained in the delimited canopies, image features describing individual trees were calculated.</p> <p>Considerable changes in tree mortality were observed. The number of dead trees at the beginning of the study period increased significantly from year 2005 to year 2009. An exceptionally dry summer in 2006 was identified as a possible reason. In the following years, the situation remained moderate, but in quantitative and regional terms, mortality was at its highest in 2017. Overall, there was an upward trend in mortality during the study period, and average annual temperatures were found correlating strongly with the number of dead trees ($r=0.73$). The classification accuracies of tree health status varied annually between 89–96%. The seven-year average accuracy was 93.6% with a kappa value of 0.88. The most important features in the classification were the features calculated from the blue channel, such as the maximum value of the channel (B_{max}), the difference between the maximum and minimum of the channel (B_{range}) and the skewness of the distribution (B_{skew}).</p> <p>The results of the thesis showed that tree mortality can be monitored using remote sensing data. Clear changes in the number of dead trees were observed using the time series and possible causes were identified. By identifying the causes behind rising mortality, the effects of climate change can also be better understood. Tree health status classification accuracies were at a good level and dead trees can be mapped from aerial imagery by semi-automatic methods. Based on the thesis, it can be rightly stated that changes in tree mortality can be observed with aerial imagery time series. In addition, the semi-automatic identification of dead trees from aerial imagery can be said to be accurate enough for large-scale use.</p>			
Avainsanat – Nyckelord – Keywords Tree mortality, remote sensing, climate change, tree health, machine learning			
Säilytyspaikka – Förvaringställe – Where deposited Helsinki University Library – Helda / E-thesis (theses) ethesis.helsinki.fi			
Muita tietoja – Övriga uppgifter – Additional information			

Sisällysluettelo

1. JOHDANTO	5
1.1 Työn tausta.....	5
1.2 Teoria.....	7
1.2.1 Puiden kuolleisuus.....	7
1.2.2 Ilmakuvaus	8
1.2.3 Laserkeilaus	9
1.2.4 Kaukokartoitus metsien terveydentilan seurannassa	11
1.3 Tutkimuksen tavoitteet	13
2. AINEISTO JA MENETELMÄT	14
2.1 Tutkimusalue	14
2.2 Kaukokartoitusaineistot	16
2.2.1 Ilmakuva-aineisto	16
2.2.2 Laserkeilausaineisto	16
2.3 Referenssiaineisto	16
2.4 Keskilämpötilat ja sademäärät.....	18
2.5 Puiden kuolleisuuden analysointi	19
2.6 Kaukokartoitusaineistojen käsittely	20
2.7 Sävyarvojen irrotus	22
2.8 Kasvillisuusindeksien laskenta.....	24
2.9 Puiden puoliautomaattinen luokittelu	26
2.10 Luokittelun tarkkuuden arviointi.....	28
3. TULOKSET	30
3.1 Kuolleisuuden kehitys keskuspuiston alueella tutkimusjakson aikana	30
3.2 Kuolleisuuden kehitys keskuspuiston pohjoisosassa	32
3.3 Luokittelutarkkuus terveydentilan mukaan	34
4. TULOSTEN TARKASTELU	38
4.1 Kuolleisuuden kehityksen tarkastelu	38
4.2 Puiden terveydentilan luokittelutarkkuuden tarkastelu	41
4.3 Luokittelumenetelmien erojen tarkastelu	45
5. JOHTOPÄÄTÖKSET	46
6. LÄHTEET.....	49
LIITTEET	60

1. JOHDANTO

1.1 Työn tausta

Puut kohtaavat yhä moninaisempia terveysuhkia, kuten uusille alueille leviäviä tuhonaiheuttajia, muuttuvien elinolosuhteiden aiheuttamaa toistuvaa stressiä sekä mahdollisesti uusia vielä tuntemattomia uhkia. Ilmastonmuutoksen kokonaisvaikutuksia puihin ja metsiin on vaikea ennustaa. Lämpenevän ilmaston seurauksena kasvukausi pitenee parantaen puuston kasvua, mutta samalla kuivuus- ja hyönteistuhot saattavat yleistyä ja myrskytuhojen riski kasvaa (Luonnonvarakeskus 2021a). Lisäksi lämpeneminen on keskimääräistä suurempaa boreaalisella vyöhykkeellä. Nousevat lämpötilat sekä epätasaisesti jakautuvat sademäärät vuodenaikojen välillä luovat olosuhteet pitkillekin kesäaikaisille kuivuusjaksoille (Kinnunen ym. 2013). Kuivuus johtaa usein puiden kuolemaan ja on etenkin kuuselle vakava uhka (Parviainen ym. 2010). Kuivuus myös aiheuttaa puille stressiä leutojen talvien tapaan, jolloin puiden puolustautumismekanismit heikentyvät altistaen ne esimerkiksi hyönteistuhoilta (Ayres ja Lombardero 2000, Marini 2012 viitattu Kinnunen ym. 2013). Lämpötilojen nousu mahdollistaa myös uusien tuhohyönteislajien leviämisen Suomeen (Parviainen ym. 2010).

Puiden kuolleisuuden seurannalla on monia hyötyjä. Sillä voidaan arvioida metsien terveydentilaa, tuottavuutta, hiilensidontaa sekä kuolleiden puiden ekologisia vaikutuksia metsäekosysteemeissä (Meddens ym. 2011, Sylvain ym. 2019). Kuollut puusto on arvokasta biodiversiteetin kannalta, ja kuolleisuuden seurannalla voidaan seurata myös lahopuun syntymistä. Puun kuolemaan johtavia syitä voivat olla muun muassa biologiset, ilmastolliset tai ihmiseen liittyvät tekijät (Sylvain ym. 2019). Seuranta voi lisätä ymmärrystä kuolemaan johtavista syistä ja mahdollisesti auttaa metsien paremmassa suojelussa. Hyönteistuhot voivat olla alueellisesti hyvin vaihtelevia, jolloin puut eivät välttämättä kuole ryhmittäin vaan varsinkin alussa yksittäin (Meddens ym. 2011). Tarkastelua tulisi-kin tehdä puutasolla ymmärtääksemme paremmin tuhojen lisäksi puiden ekologiaa, sillä yksittäiset puut voivat selvitä vakavistakin epidemioista (Meddens ym. 2011, Hart ja Veblen 2015). Puutason tarkastelu mahdollistaa myös tuhojen havaitsemisen ajoissa ensimmäisten puiden kuollessa. Alkuvaiheessa havaitut metsätuhot auttavat minimoimaan

taloudelliset tappiot talousmetsissä, ja virkistysmetsissä vaikutuksia voidaan pyrkiä rajoittamaan.

Perinteisesti kuolleiden puiden kartoitus Suomessa on ollut passiivista perustuen maastotai ilmakuvahavaintoihin. Aktiivisesti kartoitusta on tehty vain isojen metsätuhojen esimerkiksi myrskyjen jälkeen kartoituslennoin. Puihin liittyvä kuolleisuus on spatiaalisesti ja ajallisesti epäsäännöllinen prosessi, jonka seuraaminen on hankalaa perinteisin inventointimenetelmin (Sylvain ym. 2019). Kaukokartoitus mahdollistaakin puiden kuolleisuuden tehokkaamman kartoittamisen ja seurannan. Sen hyödyllisyys etenkin hyönteisten aiheuttamaa kuolleisuutta tutkittaessa on yleisesti tunnistettua (mm. Franklin ym. 2003, Skakun ym. 2003, Wulder ym. 2006, Hart ja Veblen 2015). Kaukokartoitusaineistot ovat objektiivisia, ajallisesti hyvin toistettavissa, spatiaalisesti kattavia sekä havaitsevat tarkimmillaan jopa yksittäiset puut (Meddens ym. 2011, Freeman ym. 2016). Lisäksi kaukokartoituksen on todettu olevan jopa 90 % halvempaa kuin perinteinen maastossa tehtävä aineistojen keruu (Latifi ym. 2015).

Kuolleiden puiden kartoitusta satelliitti- ja ilmakuvilta onkin tutkittu runsaasti (mm. Meddens ym. 2011, Hart ja Veblen 2015, Freeman ym. 2016, Van Gunst ym. 2016, Wang ym. 2016, Kellner ja Hubbell 2017, Sylvain ym. 2019, Bright ym. 2020). Vasta viime vuosina tutkimuksissa on osoitettu pelkkiin kaukokartoitusaineistoihin perustuvien operationaalisten menetelmien toimivuus yksittäisten kuolleiden puiden kartoituksessa laajoilla alueilla (Sylvain ym. 2019, Krzystek ym. 2020). Kuolleisuuden pitkän ajan trendien seuranta aikasarjojen avulla voi paljastaa uusia asioita pidemmän ajan tuhoista, jotka eivät ole nähtävissä yhden ajankohdan aineistoissa (Bright ym. 2020). Esimerkiksi tieto hyönteistuhon suunnasta ja voimakkuudesta sekä niiden muutoksista auttaa ymmärtämään paremmin niiden leviämistä (Hart ja Veblen 2015). Valtaosa aiemmista aikasarjoista hyödyntävistä tutkimuksista on käyttänyt satelliittikuvia niiden paremman ajallisen toistettavuuden takia. Ilmakuva-aikasarjat tarjoavatkin uusia mahdollisuuksia puiden kuolleisuuden kartoittamiseen.

1.2 Teoria

1.2.1 Puiden kuolleisuus

Puun kuoleman aiheuttaa aina jokin tapahtuma, ja usein kuolemaan voi olla useampiakin syitä (Lundquist ym. 2011). Nämä kuolemaan johtavat syyt tai tekijät voidaan jakaa abioottisiin ja bioottisiin. Abioottiset tuhonaiheuttajat ovat elottomia, näitä ovat esimerkiksi myrskyt, kuivuus sekä lumi (Luonnonvarakeskus 2021b). Myös ravinnepuutokset ja ilmansaasteet voivat aiheuttaa puun kuoleman (Lundquist ym. 2011). Bioottisten eli elävien tuhonaiheuttajien ryhmä on huomattavasti laajempi ja se voidaan jakaa sienitauteihin, hyönteisiin sekä selkärangkaisiin (Luonnonvarakeskus 2021b). Perinteisesti Suomessa yleisimpiä metsätuhoja ovat olleet lumituhot, hirvituhot, lahottajasienten tuhot sekä tuulituhot (Nuorteva 2019). Lämpenevä ilmasto suosii etenkin erilaisia kaarnakuoriaisia (Ilmastokatsaus 2019) kasvattaen niiden osuutta tulevista tuhoista. Mahdollisesti suurimman tuhopotentiaalin Suomessa tulevaisuudessa omaava hyönteinen on kirjanpainaja (*Ips typographus*).

Kirjanpainaja on merkittävin metsätuhoja aiheuttava kaarnakuoriainen Euroopassa (Christiansen ja Bakke 1988 viitattu Kärvemö ym. 2014a). Kirjanpainajan parveilu alkaa keväällä tai alkukesästä ilman lämpötilan ylittäessä 18 °C sekä maan lämpötilan kohottua yli 9 °C (Kankaanhuhta ym. 2014). Parveilun aikana kirjanpainajat tarvitsevat lisääntymismateriaalia, jolloin ne iskeytyvät kuusien runkoihin (Kankaanhuhta ym. 2014). Munissaan naaraat kaivavat emokäytävän kaarnan alle puun pituussuuntaisesti (Kankaanhuhta ym. 2014). Munista kuoriutuneet toukat sekä koteloitumisen jälkeen kuoriutuneet aikuiset syövät puun nilakerrosta heikentäen latvuksen ravinteiden saantia, jolloin puu alkaa kuivua (Kankaanhuhta ym. 2014). Kirjanpainajien kuljettamat sinistäjä sienet haittaavat myös omalta osaltaan puiden vesi- ja ravinnekuljetuksia (Kankaanhuhta ym. 2014). Kirjanpainajien ja sinistäjä sienien yhteisvaikutukset johtavat puiden kuolemaan nopeasti, useimmiten jopa kuukausissa.

Lisääntymismateriaalin määrä määrittää kirjanpainajakannan kokoa ja jos lisääntymismateriaalia esimerkiksi tuulenkaatoja on runsaasti voi kirjanpainajakanta runsastua nopeasti (Kankaanhuhta ym. 2014). Lisäksi leuto talvi ja aikainen sekä lämmin kevät luovat

edellytykset kirjanpainajan toiselle parveilulle ja näin myös toiselle sukupolvelle, mikäli myös kesä on erityisen lämmin ja pitkä (Pouttu ja Annila 2010).

Kirjanpainajakanta voidaan jakaa endeemiseen ja epideemiseen tasoon. Endeeminen taso on näistä tavanomaisempi, kun taas epideeminen taso tarkoittaa korkeaa kirjanpainajakantaa ja mahdollista uhkaa metsälle (Annila 2001). Endeemisellä tasolla kirjanpainajat lisääntyvät tuulenkaadoissa, lumi- sekä myrskytuhoissa ja heikentyneissä pystypuissa (Annila 2001, Kärvemö ja Schroeder 2010). Epidemian aikana kirjanpainajat kykenevät iskeytymään myös terveisiin puihin ja lisääntymään niissä, aiheuttaen puiden kuoleman (Annila 2001, Kärvemö ja Schroeder 2010). Annilan mukaan (2001) epidemian aikana kirjanpainajat voivat tappaa terveitä kuusia laajoillakin alueilla ja epidemia voi kestää vuosia.

Kirjanpainajat voivat levitä jopa yli 500 metrin päähän alkuperäisestä tuhosta (Wermelinger 2004), mutta valtaosa iskeytyy huomattavasti lähempänä oleviin puihin. Erään tutkimuksen mukaan jopa 90 % uusista kirjanpainajista iskeytyi läheisiin puihin, enintään sadan metrin päähän alkuperäisestä tuhosta (Wichmann ja Ravn 2001 viitattu Wermelinger 2004). Tutkimustulos mitattiin epidemian aikana, mutta siitä voidaan todeta puiden tai puuryhmien poiston olevan tärkeää, etenkin jos lähialueella on riskialttiita kuusia. Samankaltaisia tuloksia ovat saaneet Schroeder ja Lindelöw (2002) sekä Kärvemö ym. (2014b). Tulevaisuudessa ennakoivien toimenpiteiden merkitys korostuu, sillä ilmastonmuutoksen odotetaan lisäävän kirjanpainajatuhoja (Jönsson ym. 2007 viitattu Kärvemö ym. 2014a).

1.2.2 Ilmakuvaus

Ilmakuvaus on passiivista kaukokartoitusta, eli tarkastellaan tutkittavien kohteiden esimerkiksi puiden heijastamaa sähkömagneettista säteilyä, joka on peräisin auringosta (Holopainen ym. 2015). Nykyiset ilmakuvat ovat digitaalisia, mikä tarkoittaa sitä, että jokaisella pikselillä on oma sävyarvo, joka vastaa kohteen säteilyä (Holopainen ym. 2015). Näitä arvoja voidaan tulkita numeerisesti, jolloin saadaan tietoa kohteesta. Ilmakuvat eroavat toisistaan käytettyjen aallonpituuksien perusteella. Tavallinen väri-ilmakuva (RGB, Red-Green-Blue) sisältää näkyvän valon aallonpituusalueen, kun taas väärävärikuva (CIR, Color-infrared tai IRG, Infrared-Red-Green) sisältää sinisen kanavan sijasta

lähi-infrapunan aallonpituusalueen (NIR, near infrared). Väärävärikuvan eroavaisuudet värikuvaan verrattuna perustuvat lehtipuiden paljon voimakkaampaan säteilyyn lähi-infrapun alueella verrattuna havupuihin, jolloin lehtipuut voidaan tunnistaa väärävärikuvalta punertavina (Holopainen ym. 2015).

Ilmakuvien, kuten myös tavallisten kuvien tarkkuutta voidaan kuvata resoluutiolla, joka kertoo kuva-aineiston erotuskyvyn. Erotuskyky vaikuttaa merkittävästi kuvien hyödyntämiseen ja niistä saatavaan tietoon. Erotuskyky voidaan jakaa vielä spatiaaliseen, spektriseen sekä temporaaliseen resoluutioon (Holopainen ym. 2015). Spatiaalinen resoluutio kertoo pienimmän kuvalta erotettavan kohteen koon, ja mikäli kuvien pikselikokoa ei ole esikäsittelyssä muutettu, voidaan spatiaalinen resoluutio tulkita pikselikokona (Holopainen ym. 2015). Spektrinen resoluutio kertoo kyvystä tulkita säteilyn eri aallonpituuksia eli toisin sanoen kuinka kapeita spektrin alueita käytetyllä järjestelmällä voidaan mitata (Holopainen ym. 2015). Temporaalinen resoluutio kertoo puolestaan kuvauksen toistettavuudesta, joka voidaan tulkita ajaksi kahden kuva-aineiston välillä.

Ilmakuvista saadaan tuotettua ortokuvia valittuun koordinaatistoon orto-oikaisun avulla, jolloin saadaan mittatarkka ilmakuvapohjainen kartta (Holopainen ym. 2015). Orto-oikaisussa keskusprojektiossa oleva kuva muutetaan karttaprojektioon hyödyntämällä alueen maastomallia sekä kuvien orientaatiotietoja (Haggrén 2002). Näin kuvasta saadaan poistettua maaston korkeuseroista ja kuvan kallistuskulmista johtuva vaihtelu (Haggrén 2002). Orto-oikaisu tuottaa kuviin jonkun verran projektiovirheitä johtuen lähinnä alkuperäisten kuvien perspektiiveistä korostuen kuvien reunoilla (Haggrén 2002). Nämä projektiovirheet saavat esimerkiksi puut ja rakennukset näkymään paikoitellen vinossa tai vääristyneinä.

1.2.3 Laserkeilaus

Laserkeilaus on aktiivinen kaukokartoitusmenetelmä. Laserkeilaus mahdollistaa ilmakuvauksesta poiketen myös korkeustiedon saamisen kohteesta. Laserkeilan lähettää laserpulsseja, jotka kohteeseen osuessaan palaavat yhtenä tai useampana paluukaikuna takaisin (Vosselman ja Maas 2010). Kohteen ja keilaimen välinen etäisyys, voidaan mitata laserpulssin kulkuajasta (Holopainen ym. 2013). Kun lisäksi tiedetään, keilaimen tarkka paikka ja asento saadaan jokaiselle kohteelle tarkat koordinaatit (Holopainen ym. 2013).

Eri laserkeilausmenetelmät voidaan jaotella lentolaserkeilaukseen (ALS, airborne laser scanning), maastolaserkeilaukseen (TLS, terrestrial laser scanning) sekä liikkuvaan maastolaserkeilaukseen (MLS, mobile laser scanning).

Lentolaserkeilaus on laajimmin operatiivisessa käytössä oleva laserkeilausmenetelmä, ja sillä on tehty useampia koko maan kattavia laserkeilausinventointeja (Holopainen ym. 2013). Lentolaserkeilauksessa laserkeilauslaitteisto kiinnitetään esimerkiksi lentokoneeseen, helikopteriin tai miehittämättömään lennokkiin (UAV, unmanned aerial vehicle). Maastolaserkeilaus tapahtuu paikallaan, laserkeilaimen ollessa sijoitettuna jalustalle, kuten esimerkiksi kolmijalalle. Liikkuvassa maastolaserkeilauksessa keilain asetetaan ajoneuvon, kuten auton tai mönkijän kyytiin. Liikkuva maastolaserkeilaus voidaan toteuttaa jatkuvalla liikkeellä tai välillä pysähdellen (Holopainen ym. 2013). Mainittavimpia eroja lentolaserkeilauksen ja maastolaserkeilauksen välillä ovat pistetiheys sekä sen tasaisuus. Lentolaserkeilauksessa pistetiheydet ovat maastolaserkeilausta pienempiä, johtuen keilausetaisyysistä, mutta samasta syystä pistetiheydet ovat lentolaserkeilausaineistoissa tasaisempia (Holopainen ym. 2013).

Metsävarojen inventointi lentolaserkeilauksella voidaan jakaa nykyisellään yksinpuintulkintaan (ITD, Individual tree detection, Hyypä ja Inkinen 1999) sekä aluepohjaiseen menetelmään (ABA, Area-based approach, Naesset 1997). Aluepohjainen inventointimenetelmä toimii myös harvapulssisella laserkeilausaineistolla, mutta vaatii aina tuekseen maastoreferenssiaineiston (Holopainen ym. 2013). Vastaavasti taas yksittäisten puiden tulkinta ei ole mahdollista harvapulssisesta keilausaineistosta. Tiheämpipulssiset laserkeilausaineistot ovat kalliimpia, mutta yksinpuintulkinta mahdollistaa puustotietojen tuottamisen ilman maastomittauksia (Holopainen ym. 2013). Käytännössä laajojen metsäalueiden inventointi on perustunut tähän mennessä aluepohjaiseen menetelmään yksinkertaisesti siitä syystä, että isojen laserkeilausalueiden aineistot ovat olleet liian harvapulssisia yksinpuintulkintaan.

Laserpulssien osuessa puiden latvustoon se voi heijastua takaisin kokonaan tai läpäistä osittain latvuston (Vosselman ja Maas 2010). Latvuston läpäisyprosentti vaihtelee riippuen metsistä, mutta myös laserkeilaimen ominaisuuksista (Holopainen ym. 2013). Latvuston läpäisevät laserpulssit antavat tietoa esimerkiksi maanpinnan muodoista. Eri koh-

teista heijastuneet paluukaiut voidaan luokitella erikseen, jolloin vain maahan asti osuneista laserpisteistä voidaan muodostaa maanpintamalli (DTM, Digital Terrain Model). Vastaavasti vain latvustoon osuneista pisteistä saadaan muodostettua pintamalli (DSM, Digital Surface Model). Vähentämällä maanpinnankorkeus pintamallista saadaan lopputulokseksi latvuston korkeusmalli (CHM, Canopy Height Model). Latvuston korkeusmalli on useimmiten yksinpuintulkinnan lähtökohta (Holopainen ym. 2013), kuten myös tässä tutkielmassa.

1.2.4 Kaukokartoitus metsien terveydentilan seurannassa

Kaukokartoitus on ollut pitkään käytössä metsien havainnoinnissa. Ilmakuvia on hyödynnetty metsien terveyden seuraamiseen jo 1930-luvulta lähtien (Riley 1989). Myöhemmin mukaan ovat tulleet satelliittikuvat sekä laserkeilaus. Pystyssä olevat puut voidaan jakaa terveydentilan mukaan terveisiin, sairaisiin ja kuolleisiin puihin. Sairaisiin puihin luetaan yleisimmin värioireiset sekä harsuuntuneet puut. Puiden terveydentila ei kuitenkaan ole yksiselitteinen käsite, sillä se muuttuu jatkuvasti. Sairaat tai heikentyneet puut voivat myös palautua täysin terveiksi ajan kuluessa. Kaukokartoituksessa haastavaa onkin ajoitus, jotta voidaan havainnoida joskus nopeastikin tapahtuvia muutoksia puiden terveydentilassa.

Tutkimuksissa tarkastelu on vaihdellut joko elävien ja kuolleiden puiden, terveiden ja sairaiden puiden tai kaikkien näiden välillä. Etenkin ilmakuvilta neulaskadon tunnistaminen on hankalaa, varsinkin lievissä tapauksissa, mutta kuolleiden puiden havaitseminen on selvästi tarkempaa (Haara ja Nevalainen 2002). Ilma- ja satelliittikuvilla lähi-infrapun aallonpituusalueen hyödyllisyys kuolleiden puiden kartoituksessa on tunnistettu jo 1970-luvulla, sillä neulasten vähentyessä lähi-infrapunasäteily vähenee merkittävästi (Colwell 1970, Boehnel 1976, Vogelmann and Rock 1986, Brockhaus ym. 1989, Haara ja Nevalainen 2002, Heikkilä ym. 2002).

Aiemmat puiden terveydentilaa tarkastelevat tutkimukset ovat käyttäneet kaukokartoitusaineistoina ilmakuvia (mm. Haara ja Nevalainen 2002, Heikkilä ym. 2002, Meddens ym. 2011, Freeman ym. 2016, Sylvain ym. 2019), satelliittikuvia (mm. Meddens ym. 2013, Long ja Lawrence 2016, Van Gunst ym. 2016, Wang ym. 2016, Kellner ja Hubbell 2017,

Bright ym. 2020), näiden yhdistelmiä (mm. Hart ja Veblen 2015) sekä laserkeilausaineistoja (mm. Yao ym. 2012, Polewski ym. 2015). Lisäksi esimerkiksi Polewski (2017) sekä Krzystek ym. (2020) hyödynsivät tutkimuksissaan ilmakuvia ja laserkeilausaineistoja. Kaikilla edellä mainituista aineistoilla on kyetty tunnistamaan kuolleita puita yli 90 % tarkkuudella.

Sylvain ym. (2019) käyttivät luokittelussaan neuroverkkoa, ja erittäin laajaa 292 000 puun luokitteluaineistoa. Kaukokartoitusaineistona toimi nelikanavainen (RGBI, Red-Green-Blue-Infrared) ortoilmakuva, jonka spatiaalinen resoluutio oli 20 cm. Kokenut kuvatulkitsija luokitteli kaikki yli seitsemän metrin puut valituilta yhden hehtaarin polygoneilta puulajin (havupuu/lehtipuu) ja terveydentilan mukaan (elävä/kuollut). Puun sijainti merkittiin keskelle näkyvissä olevaa puun massaa. Puun latvusta ei segmentoitu laserkeilausaineistosta, vaan jokaisen puun ympärille luotiin ikkuna, joka rajasi latvuksen. Ikkunan koko määritettiin vastaamaan keskimääräisen latvuksen pinta-alaa. Muutamaa erikoista ikkunaa kokeiltiin, mutta ikkunan koon ei havaittu vaikuttavan juurikaan kuolleiden puiden luokittelutarkkuuteen. Tutkimuksessa käytetty menetelmä latvuksen rajaamiseen ei ole yhtä hyvin yleistettävissä vaihtelevan kokoisiin metsiköihin verrattuna latvus-segmenttimenetelmään.

Krzystek ym. (2020) käyttivät luokitteluissa kahta eri koneoppimismenetelmää (Random Forest ja logistinen regressio) ja kaukokartoitusaineistona toimivat sekä ortoilmakuvat (spatiaalinen resoluutio 17 cm ja 9,5 cm) että laserkeilausaineisto. Kuolleiden puiden luokittelussa käytettiin 9 704 puun visuaalisesti kerättyä aineistoa sekä kuva- ja laserpiirteitä. Kuolleet puut luokiteltiin kuolleiksi vain, jos luokan todennäköisyys oli yli 90 %. Lisäksi he luokittelivat kuolleita puita, joilla ei ole enää latvusta ollen käytännössä pelkkiä runkoja. Näitä runkopuita ei pystytä havaitsemaan ilmakuvilta, joten he hyödynsivät niiden tunnistamisessa laserpiirteitä. Molemmissa kuolleiden puiden luokitteluissa tarkkuudet olivat yli 90 %. Osittain tuloksia heikensivät lehdettömät puut, jotka sekoittuivat luokittelussa virheellisesti kuolleiksi puiksi.

Viime vuosina tutkimukset ovat keskittyneet entistä enemmän mahdollisimman aikaiseen puiden heikentymisen havaitsemiseen. Tavoitteena on havaita sairaat puut heti alkuvaiheessa jo ennen visuaalisten oireiden alkamista, jolloin voidaan reagoida entistä parem-

min. Näsi ym. (2020) tutkivat useiden eri häiriöiden aiheuttaman heikentymisen havaitsemista dronen avulla kerätyistä multi- ja hyperspektriaineistoista. He saivat kansainvälisestikin ensimmäisiä lupaavia tuloksia kirjanpajan, juurikäävän sekä ravinnepuutosten aiheuttamista puiden heikentymisistä. Vaikka tulokset olivatkin lupaavia, totesivat he menetelmien vaativan vielä kehitystä. Myös monikanavalaserkeilauksen mahdollisuuksia on tutkittu, ja sillä kyettiin havaitsemaan neulasten laskenut vesipitoisuus heti kirjanpajan iskeytymisen jälkeen ennen mitään ulkoisesti havaittavaa oiretta (Junttila 2019).

Kaukokartoitukseen perustuva tarkastelu ei ole täysin ongelmaton. Sylvain ym. (2019) totesivatkin pelkkiin kaukokartoitusaineistoihin perustuvien laajojen alueiden kuolleiden puiden kartoituksen olevan vielä tällä hetkellä lähes olematon. Heidän mukaansa tämän selittää osaltaan kaukokartoitusaineistojen ominaisuudet, kuten isoja alueita kartoitettaessa syntyvät epätasaisuudet aineistoon. Nämä epätasaisuudet aineistojen laadussa voivat johtua esimerkiksi eroavaisuuksista aineistojen hankinta-ajankohdissa sekä ilmakehän olosuhteista ja muutoksista kartoitusalueella (Sylvain ym. 2019).

Ilma- tai satelliittikuvien sävyarvojen tulkinta voi olla välillä epävarmaa, koska valo heijastuu puista eri kulmista, jolloin samankaltaiset puut voivat saada erilaisia sävyarvoja (Haara ja Nevalainen 2002, Heikkilä ym. 2002). Tämä voi korostua riittävän tarkoilla kuvilla myös latvuksen sisällä osan latvuksesta ollessa varjon peitossa (Meddens ym. 2011). Varjot vaikeuttavat muutenkin kuvien tulkintaa jyrkkien sävyerojen takia. Kuolleilla puilla kontrasti latvuksen ja maan välillä vähenee neulasten puutteen takia vaikeuttaen latvusten reunojen tunnistamista (Howard 1991 viitattu Haara ja Nevalainen 2002). Kuolleet puut voivatkin sekoittua kuvalla paljaan maan pikseleihin (Meddens ym. 2011).

1.3 Tutkimuksen tavoitteet

Tämän tutkielman tarkoituksena oli selvittää, miten kaukokartoitusaineistoja voidaan hyödyntää puiden kuolleisuuden seurannassa. Monivuotisista ilmakehän kuva-aineistoista ke-
rättiin laaja puuaineisto. Aineisto sisälsi terveitä, sairaita sekä kuolleita puita. Sairaista ja kuolleista puista luotiin jokaiselle vuodelle lämpökartta puiden sijaintien mukaan. Yhdistämällä vuosien kuvat luotiin aikasarja, josta tarkasteltiin kuolleisuuden määrällistä ja

alueellista vaihtelua. Kerättyä aineistoa hyödynnettiin myös puoliautomaattisissa luokitteluissa, jotka toteutettiin neljällä eri koneoppimismenetelmällä. Tavoitteena oli saada selville, millä tarkkuudella sairaita ja kuolleita puita voidaan tunnistaa ilmakuvilta.

Tutkielman tutkimuskysymyksiä olivat:

1. Miten puiden kuolleisuus on muuttunut määrällisesti ja alueellisesti tutkimusjakson aikana?
2. Onko puiden kuolleisuus lisääntynyt tutkimusjakson aikana?
3. Millä tarkkuudella kuolleiden puiden luokittelu onnistuu ilmakuvilta?
4. Miten eri koneoppimismenetelmien luokittelutarkkuudet eroavat toisistaan?

2. AINEISTO JA MENETELMÄT

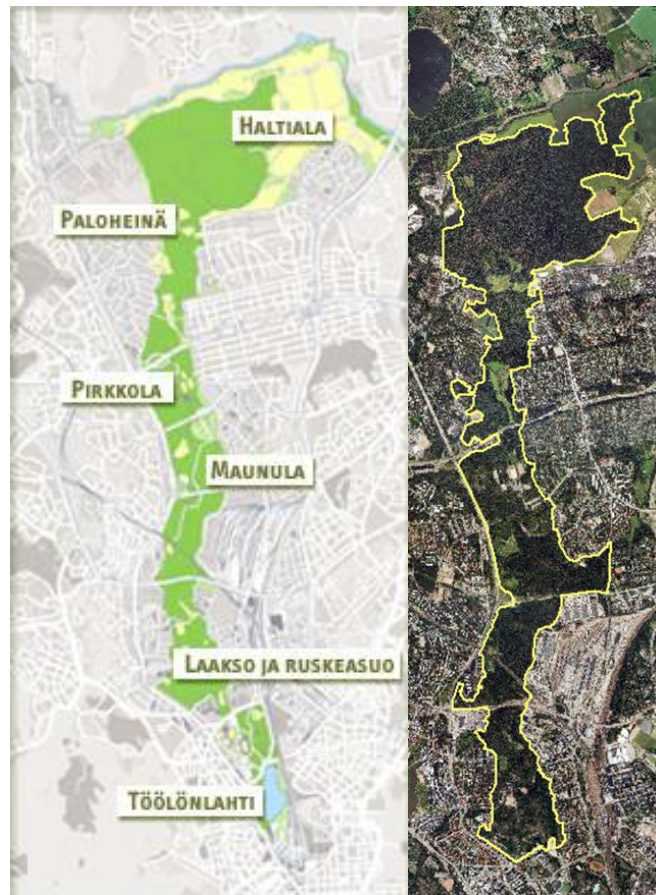
2.1 Tutkimusalue

Tutkimusalueena tutkielmassa toimi Helsingin keskuspuisto (kuva 1). Keskuspuisto on noin 1 000 hehtaarin laajuinen puistoalue, josta metsiä on noin 700 hehtaaria (Helsingin kaupunki 2020, My Helsinki 2020). Alue alkaa aivan kaupungin ytimestä Töölönlahdesta ja ulottuu aina Vantaan rajalle asti (My Helsinki 2020), joten kävijämäärät ovat suuria puiston alueella. Vuonna 2014 vuosittaisten vierailujen määräksi arvioitiin noin kaksi miljoonaa (Forest.fi 2014), mikä on enemmän kuin kaikkien kansallispuistojen vuosittaiset kävijämäärät yhteensä (Retkipaikka 2016). Puiston alueella onkin kattavasti kulkureittejä ja sen virkistyskäyttö on laajaa.

Keskuspuiston alue sekä sen puisto ovat hyvin monimuotoisia. Alue on suurilta osin luonnontilainen sekä vähän hoidettu ja sisältää muun muassa runsaasti varttuneita kuusi-valtaisia metsiä (Helsingin kaupunki 2020). Luonnontilaisuuden takia keskuspuiston puisto on pääosin luonnollisesti uudistunutta ja alueelta löytyykin hyvin erikokoista puustoa. Lajistoltaan alue on erittäin monipuolinen, vaikkakin valtaosa puustosta koostuu pääpuulajeistamme männystä, kuusesta ja koivusta. Merkittävä lehtipuu alueella on

haapa, mutta myös harvinaisempia lajeja löytyy, kuten Maunulasta löytyvä pähkinäpensaslehto osoittaa (Citynature 2020). Muita merkittäviä kohteita ovat muun muassa luonnonsuojelualueet, kuten Haltialan aarnialue, lehtoalueet sekä Niskalan arboretum eli puulajipuisto (My Helsinki 2020). Keskuspuiston alueella on kaikkiaan viisi luonnonsuojelualuetta ja ainakin kaksi muuta on vielä tulossa (Citynature 2020).

Tutkimusalueeksi valittiin keskuspuisto alueella esiintyvien kaarnakuoriaistuhojen takia. Yksittäisiä kuolleita puita löytyy koko puiston alueelta, ja lisäksi esimerkiksi Paloheinän ja Maunulan alueilla on suurempia tuhokeskittymiä. Nämä tuhokeskittymät koostuvat useista kymmenistä järeistä kuolleista kuusista. Alueen luonteen sekä osittain kaarnakuoriaisten takia puiston alueelta löytyy pystyyn kuolleiden puiden lisäksi kaatuneita ja katkenneita puita. Tässä tutkielmassa keskityttiin vain pystyssä oleviin kuolleisiin puihin eikä kaatuneita puita huomioitu. Lisäksi pystyssä olevista kuolleista puista voitiin tarkastella vain niitä, joissa oli vielä latvus, sillä latvuksettomat kuolleet puut ovat käytännössä näkymättömiä ilmakuvissa.



Kuva 1. Vasemmalla Helsingin keskuspuiston osa-alueet (Helsingin kaupunki 2020). Oikealla tutkimusalueen tarkka raja.

2.2 Kaukokartoitusaineistot

2.2.1 Ilmakuva-aineisto

Ilmakuva-aineistona käytettiin Helsingin kaupungin ortoilmakuvia. Ortoilmakuva on maan pintamallin avulla laskennallisesti tuotettu mittatarkka ilmakuva, joka kuvaa maanpinnan kohtisuoraan ylöspäin. Helsinki kuvataan nykyään ilmasta vuosittain ja säännölliset ilmakuvausaukset aloitettiin jo 1930-luvulla, tuolloin kuvausväli tosin oli harvempi (Helsingin paikkatietohakemisto 2021). Vuosittainen kuvausajankohta vaihtelee alkuvästä loppukesään, joten aineistoksi valikoitiin vain lehdelliseen aikaan kuvatut kuvat. Lehdettömään aikaan kuvatuissa kuvissa kuolleiden puiden tunnistaminen olisi heikompaa, sillä niitä on vaikea erottaa lehtipuista.

Kaikki tai lähes kaikki parillisina vuosina kuvatut ilmakuvat oli kuvattu lehdettömään aikaan, joten niitä ei voitu hyödyntää tutkielmassa. Lopulliset tutkielmaan valitut ortokuvat olivat vuosilta: 2005, 2009, 2011, 2013, 2015, 2017 sekä 2019. Vastaavassa järjestyksessä kuvien spatiaaliset resoluutiot senttimetreinä olivat: 20, 25, 20, 20, 10, 8, 20. Kuva-aineisto koostui näin ollen 7 vuoden kuvista ja kattoi 15 vuoden ajanjakson, mikä oli riittävän pitkä havainnoimaan myös isompia trendejä. Kaikki käytetyt kuvat olivat kolmikanavaisia värikuvia (RGB, Red-Green-Blue).

2.2.2 Laserkeilausaineisto

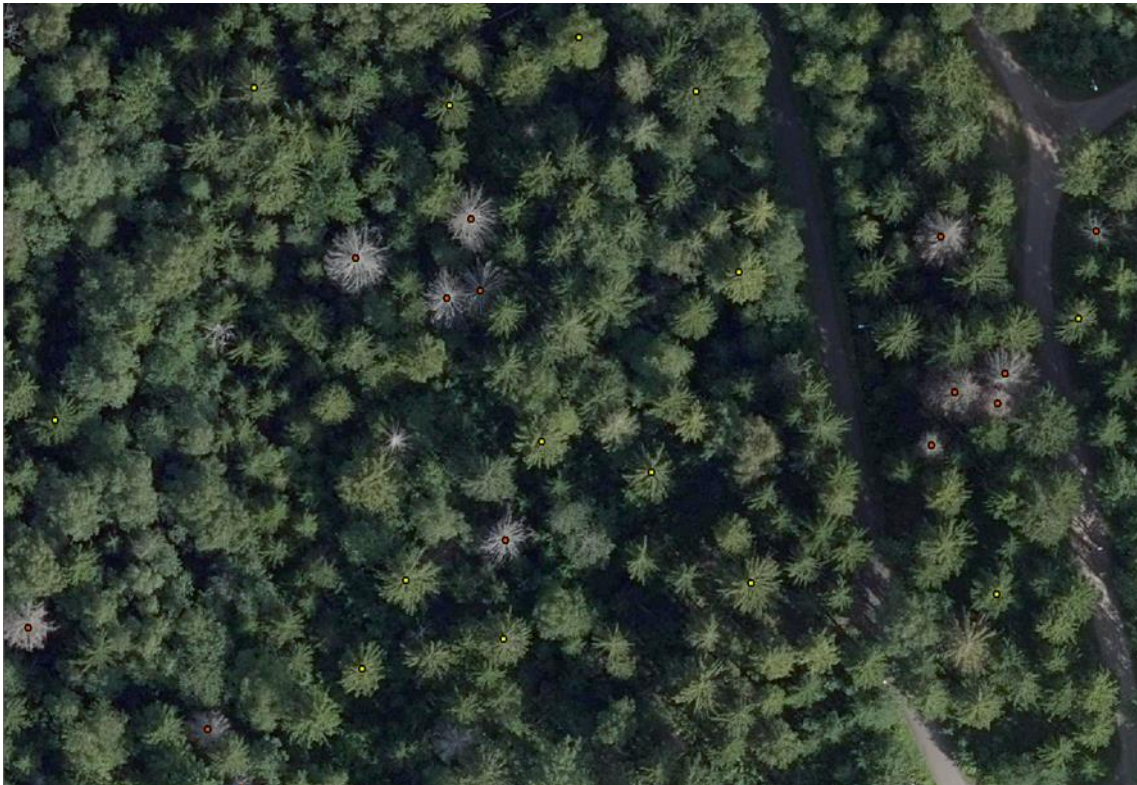
Laserkeilausaineistona käytettiin Helsingin kaupungin vuonna 2015 keräämää lentolaserkeilausaineistoa (ALS, airborne laserscanning). Käytetyssä aineistossa pistetiheys oli noin 34 pistettä/m². Laserkeilausaineisto oli ETRS-GK25 -koordinaatistossa ja N2000 -korkeusjärjestelmässä.

2.3 Referenssiaineisto

Referenssiaineisto kerättiin käsin visuaalisena kuvatulokintana ortoilmakuvilta. Jokaisen vuoden kuva käytiin järjestelmällisesti läpi koko keskuspuiston alueella tavoitteena kerätä aineistoon terveitä, kuolleita sekä sairaita niin kutsuttuja ”red attack” puita. Nämä

sairaavat puut ovat kuolemassa olevia puita, jotka näkyvät kuvilla punertavina. Muissa vastaavissa tutkimuksissa on myös pyritty tunnistamaan sairaita puita kuolleiden puiden lisäksi (mm. Meddens ym. 2011, Hart ja Veblen 2015). Kuolleista puista ei kerätty kaikkia, vaan latvuksen halkaisijaltaan noin 3–4-metriset ja sitä isommat puut otettiin mukaan aineistoon. Jokaista kuollutta puuta kohden valittiin yksi terve puu. Lisäksi kerättiin kaikki löydettyt sairaat puut, jotka täyttivät samat kriteerit, kuin kuolleilla puilla. Jokaiselle puulle merkittiin puulaji havupuun tai lehtipuun tarkkuudella.

Yksittäisen puun sijainti määritettiin puun latvaan, jos puu näkyi kuvalla suoraan ylöspäin (Kuva 2). Vinossa olevien puiden kohdalla sijainti määritettiin puun massan keskikohtaan samaan tapaan kuin Sylvain ym. (2019) omassa tutkimuksessaan. Näin puulle määritetty sijainti oli todennäköisesti lähempänä todellista sijaintia ja puun sijainnin keskipisteen ollessa mahdollisimman keskellä puuta kuvaavia pikseleitä puiden sävyarvojen irrotus onnistui paremmin.



Kuva 2. Referenssiaineiston puita, joilla keltainen piste kuvaa tervettä puuta ja punainen kuollutta.

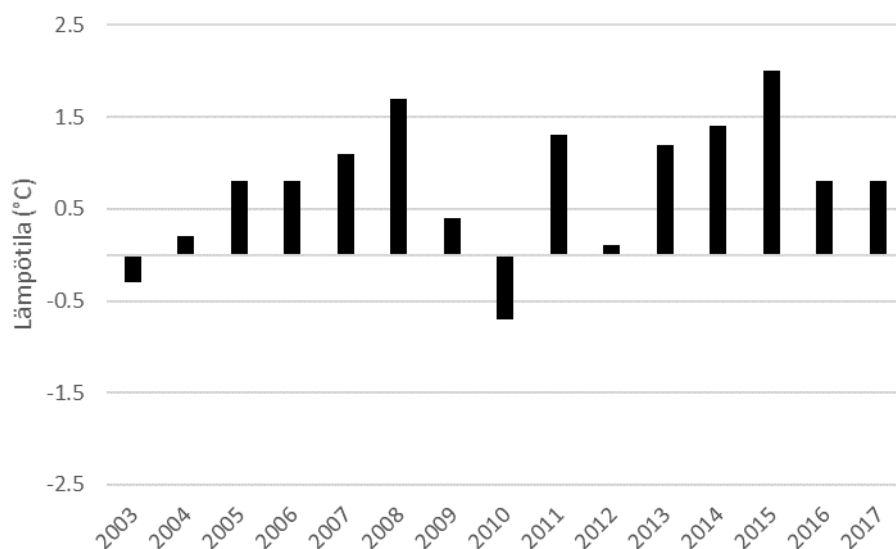
Taulukossa 1 näkyy eri vuosilta kerätty aineisto. Kokonaispuumääräksi muodostui 14 212 puuta. Terveitä ja kuolleita puita oli täsmälleen sama määrä (6 789), sairaiden puiden määrä oli huomattavasti pienempi (634).

Taulukko 1. Visuaalisella kuvatulkinalla kerätyn referenssiaineiston lukumäärät vuosittain ja luokittain.

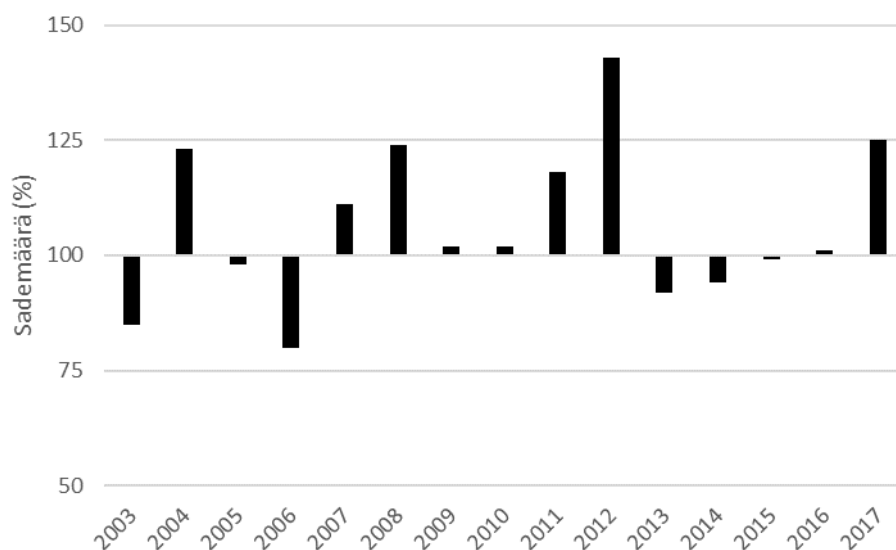
Vuosi	Yhteensä	Terveet	Sairaat	Kuolleet	Havupuut	Lehtipuut	Kuolleista lehtipuita
2005	226	113	0	113	216	10	0
2009	2736	1339	58	1339	2534	202	6
2011	2265	1035	195	1035	2006	259	12
2013	2303	1128	47	1128	1904	399	21
2015	2579	1212	155	1212	2166	413	83
2017	2818	1400	18	1400	2394	424	74
2019	1285	562	161	562	1136	149	16
Yhteensä	14212	6789	634	6789	12356	1856	212

2.4 Keskilämpötilat ja sademäärät

Tutkielmassa hyödynnettiin Ilmatieteen laitoksen vuodesta 1961 tilastoimia säähavain-toja Helsingistä. Havaintoja löytyi sekä vuosittaisesta keskilämpötilasta että sademää-rästä. Tarkastelujaksoksi valittiin 2003–2017. Kuvassa 3 näkyy vuosittaisen keskilämpö-tilan poikkeama vuosien 1981–2010 keskiarvosta, joka on 5,7 °C. Kuvassa 4 näkyy vas-taavasti vuosittainen sademäärä prosentteina vuosien 1981–2010 keskiarvosta, joka on 641 mm.



Kuva 3. Keskilämpötilan poikkeama 1981–2010 keskiarvosta (5,7 °C) Helsingissä (Ilmatieteen laitos 2021).



Kuva 4. Sademäärä prosentteina 1981–2010 keskiarvosta (641 mm) Helsingissä (Ilmatieteen laitos 2021).

2.5 Puiden kuolleisuuden analysointi

Referenssiaineiston kuolleiden puiden sijainneista muodostettiin kuolleisuutta kuvastava lämpökartta jokaisesta vuodesta. Yksittäisten vuosien lämpökartat yhdistettiin aikasarjaksi, jolla voitiin havainnoida kuolleisuuden alueellisia muutoksia. Luotuihin lämpökart-

toihin lisättiin myös sairaiden puiden sijainnit, jolloin saatiin luotua toinen aikasarja. Lisäksi laskettiin korrelaatioita säähavaintojen sekä kuolleiden puiden määrien välillä. Tutkielman tarkasteluvuosille (Taulukko 1) laskettiin kahden edellisen vuoden keskilämpötilojen ja sademäärien keskiarvot. Poikkeuksena vuosi 2009, jolle laskettiin keskiarvot edeltäviltä neljältä vuodelta. Laskettujen keskiarvojen sekä kuolleiden puiden määrien korrelaatioilla arvioitiin säämuuttujien vaikutusta puiden kuolleisuuteen.

2.6 Kaukokartoitusaineistojen käsittely

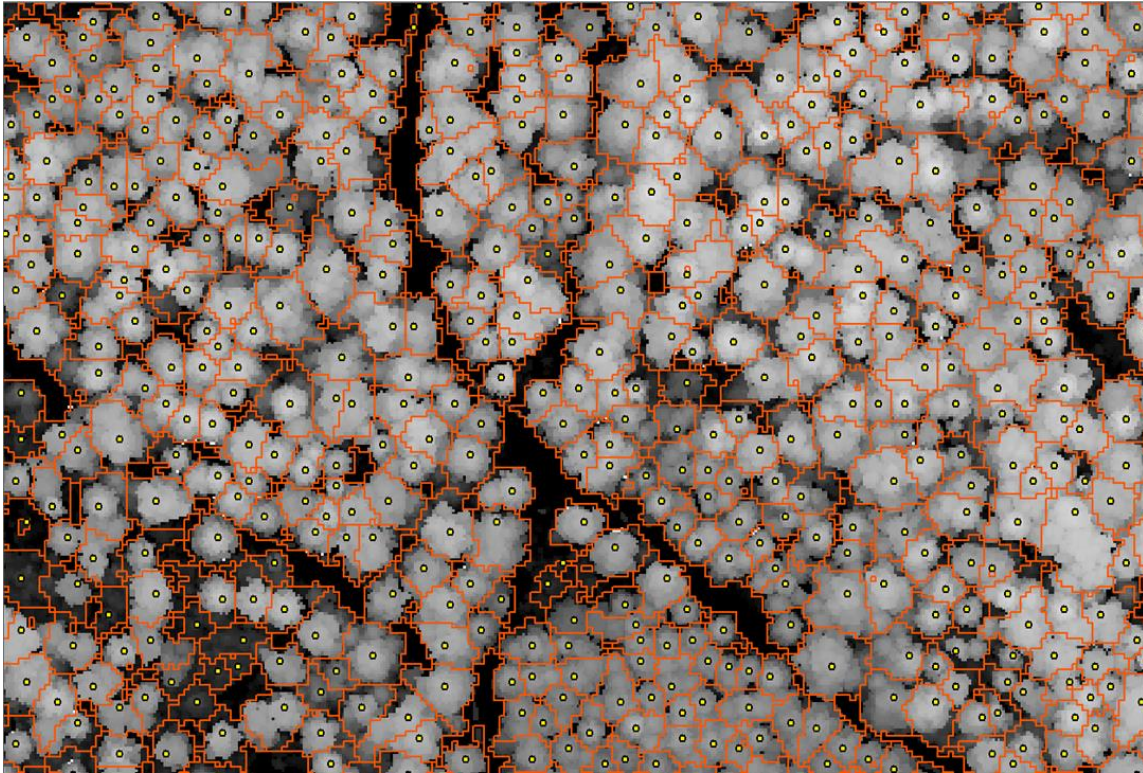
Jokaisen vuoden ortoilmakuvista muodostettiin ortomosaiikki, jotta kuvat pystyttiin georeferoimaan. Georeferointi oli tärkeää, sillä ortoilmakuvien tuli olla mahdollisimman tarkasti kohdikkain yksittäisiä puita tarkastellessa. Georeferoinnissa käytettiin CHM:n sijainteja todellisina sijainteina, joihin kuvat kohdistettiin. Kuvissa esiintyi projektiovirheiden aiheuttamia vinoumia, jotka vääristivät kohteiden sijainteja, ja hankaloittivat hyvien kiintopisteiden löytymistä. Jokaisen vuoden kuvan kohdistamiseen käytettiin kahdeksaa kiintopistettä, jotka pysyivät koko ajan samoina. Hyviksi kiintopisteiksi osoittautuivat rakennusten kulmat, sillä projektiovirheiden aiheuttamat vinoumat eivät vaikuttaneet rakennusten alakulmiin, jos ne olivat näkyvissä. Vinoumat ja vääristymät näkyivät kuvissa rakennusten yläosissa, mutta niiden alanurkat olivat aina todellisessa sijainnissa, ja tätä hyödyntämällä kuvat saatiin georeferoitua tarkasti kohdikkain CHM:n kanssa (Kuva 5). Näin ollen virheettömät kohteet saatiin täysin oikeille paikoille, kun taas edellä mainittujen virheiden takia puiden tyvi oli oikeassa kohdassa, mutta latva saattoi olla sivussa.



Kuva 5. Vasemmalla ortoilmakuvalta määritetyt rakennuksen nurkkien sijainnit näkyvät keltaisina pisteinä. Oikealla samojen pisteiden sijainnit verrattuna CHM:n todelliseen sijaintiin. Ylänurkan sijainti on neljä metriä sivussa.

Laserpisteiden muodostamaa pistepilveä käsiteltiin R-laskentaohjelman (R Core Team 2021) ”lidR” -paketilla (Roussel ym. 2021). Aineistosta luettiin vain pisteiden koordinaatit (x, y, z), luokittelu (classification), sekä paluukaiun järjestys (return). Tämän jälkeen pistepilvi normalisoitiin maaston pintamallilla (DTM, digital terrain model), joka kuvaa maanpintaa. Normalisoinnissa jokaisen laserpisteen korkeushavainnosta vähennetään maaston korkeusvaihtelut ja näin jäljelle jäävät pisteiden korkeudet kuvaavat suoraan korkeutta maanpinnasta. Normalisoidusta pistepilvestä laskettiin latvuston korkeusmalli (CHM, canopy height model) yhden metrin resoluutiolla. Ennen CHM:n laskemista aineisto rajattiin vain alle 40 metrin korkeudella oleviin pisteisiin, etteivät korkeammalla olevat mahdolliset virrehavainnot vaikuttaisi lopputulokseen.

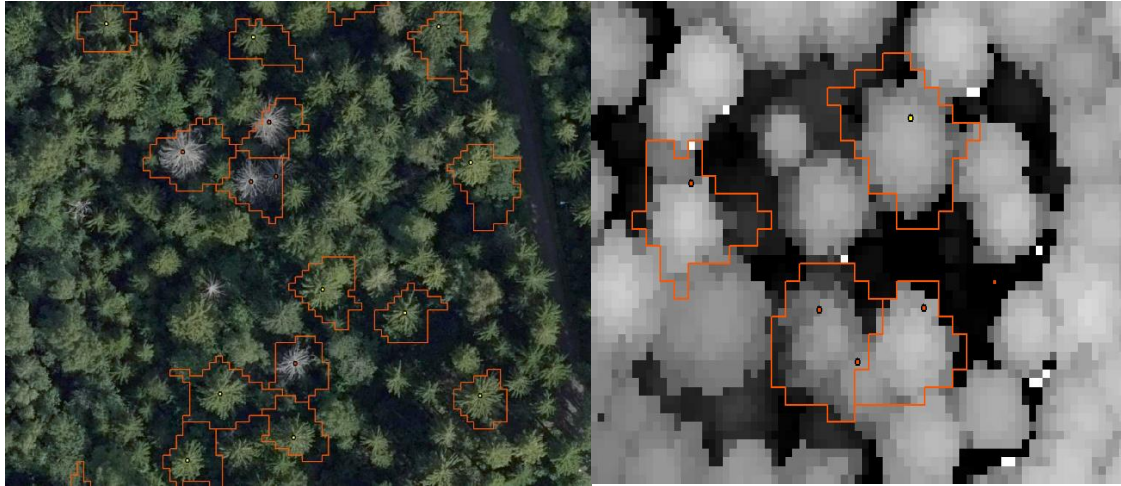
Yksinpuintulkinta toteutettiin kahdessa vaiheessa. Ensiksi CHM:ltä etsittiin yksittäisten puiden latvapisteet, jonka jälkeen rajattiin niiden latvukset. Puiden latvat etsittiin local maxima -algoritmeilla (Popescu ym. 2002, Popescu ja Wynne 2004). Algoritmi perustuu paikallisten huippuarvojen etsintään ja näiden paikallisten huippujen oletetaan olevan puiden latvoja. Korkeuden minimiarvona käytettiin neljää metriä, jolloin algoritmi ei etsi latvoja eikä tunnista puita kyseisen korkeuden alapuolelta. Löydettyjä latvapisteitä hyödyntäen segmentointiin puiden latvukset. Segmentointi toteutettiin watershed -algoritmeilla (Meyer ja Beucher 1990, Hyypä ja Inkinen 1999). Watershed tai valuma-alue-segmentointi perustuu puun latvuksen paikalliseen topografiaan, jossa veden alaspäin valumisen simuloinnilla löydetään latvuksen reunat. Lopputuloksena koko tutkimusalueesta saatiin luotua latvuskartta. Kuvassa 6 näkyy laskettu CHM sekä siitä löydettyjen puiden latvapisteet keltaisilla pisteillä ja latvusten rajat punaisina.



Kuva 6. Latvuston korkeusmalli, jossa keltaiset pisteet kuvaavat yksinpuintulkinnalla löydettyjä puiden latvapisteitä. Latvapisteiden ympärille muodostuneet latvussegmentit näkyvät punaisella.

2.7 Sävyarvojen irrotus

Referenssiaineiston puille yhdistettiin niiden sijaintien mukaisesti kohdalla oleva CHM:ltä saatu puun latvussegmentti (Kuva 7). Latvussegmentti rajaa puun latvuksen ja näin ollen ne määrittivät, mitkä pikselit kuuluivat yksittäiseen puuhun. Puiden kuvilta määritetyt sijainnit vaihtelivat todellisista sijainneista, joten kaikissa tapauksissa puille ei saatu yhdistettyä oikeaa latvussegmenttiä. Näissä tapauksissa puulle määritetty sijainti oli viereisen segmentin sisällä, jolloin puu sai viereisen puun latvussegmentin. Virheellisen latvussegmentin takia puulle rajatut pikselit saattoivat sisältää osittain tai jopa lähes kokonaan puuhun kuulumattomia pikseleitä. Latvussegmenttien yhdistämisvaiheessa aineisto myös pieneni hieman, sillä joidenkin puiden sijainti osui luodun latvuskartan tyhjään kohtaan. Nämä tyhjät kohdat olivat useimmiten aukkoja latvustossa tai taimikkoja, joiden korkeus jäi alle yksinpuintulkinnan minimikorkeuden.



Kuva 7. Vasemmalla referenssiaineiston puita sekä niille yhdistettyjä latvussegmenttejä. Oikealla näkyy yksi kuollut puu, jonka sijainti osui CHM:n puuttomaan kohtaan, jolloin latvussegmenttiä ei löytynyt.

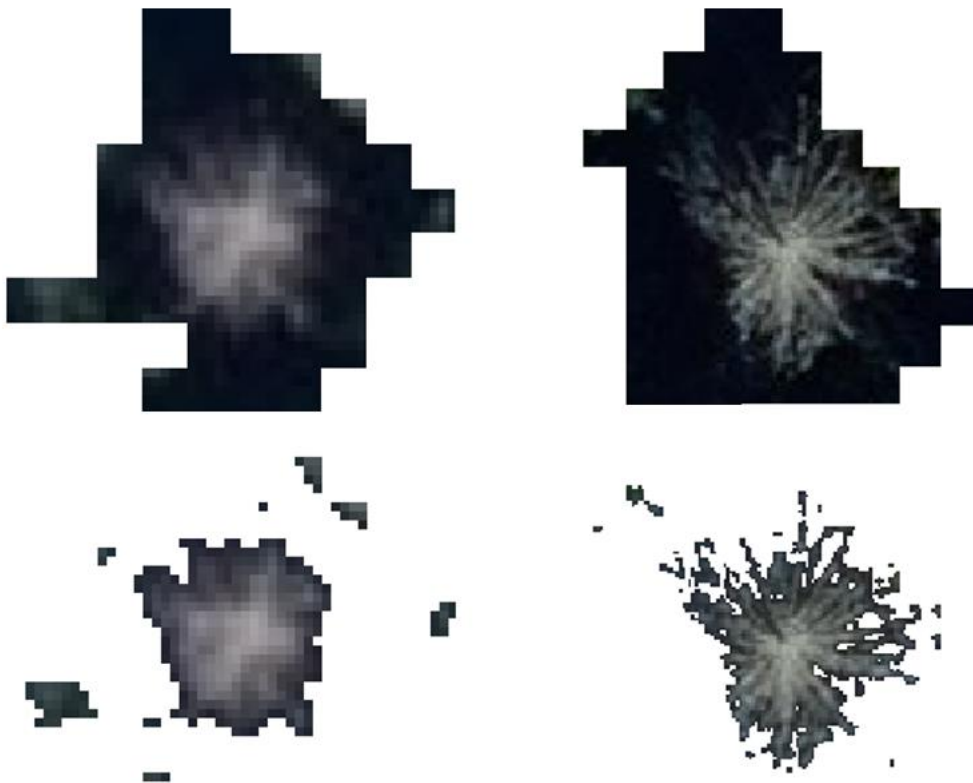
Taulukossa 2 näkyy lopullinen luokitteluaineisto. Kaikkiaan 1 165 puulle ei löytynyt latvussegmenttiä, jolloin luokitteluaineiston koko oli 13 047 puuta. Kuolleiden puiden määrä pieneni merkittävästi terveitä enemmän luoden pienen epätasapainon luokkien välille. Vuosista eniten pieneni 2009 vuoden aineisto 374 puulla.

Taulukko 2. Puoliautomaattisiin luokitteluihin käytetty aineisto vuosittain ja luokittain.

Vuosi	Yhteensä	Terveet	Sairaat	Kuolleet	Havupuut	Lehtipuut	Kuolleista lehtipuita	Ero referenssi-aineistoon
2005	192	102	0	90	182	10	0	-34
2009	2362	1298	33	1031	2163	199	3	-374
2011	2012	1007	141	864	1759	253	12	-253
2013	2084	1089	35	960	1694	390	20	-219
2015	2460	1204	133	1123	2059	401	74	-119
2017	2701	1383	17	1301	2281	420	73	-117
2019	1236	554	149	533	1088	148	16	-49
Yhteensä	13047	6637	508	5902	11226	1821	198	-1165

Latvussegmenttien yhdistämisen jälkeen jokaiselle puulle irrotettiin sitä kuvaavat pikselit segmenttien rajausten perusteella. Näiden pikselien sävyarvoista voitiin laskea puille niitä kuvaavia kuvapiirteitä. Ennen kuvapiirteiden laskentaa varjoiset pikselit rajattiin pois, sillä ne eivät kuvaa puuta ja niiden sävyarvot vääristävät laskettuja piirteitä (Kuva 8). Varjojen rajausta perustui punaisen kanavan sävyarvoihin. Pikselin punaisen kanavan sävyarvon jäädessä alle määritetyn minimiarvon se jätettiin pois laskennasta. Bright ym.

(2020) käyttivät punaisen kanavan prosenttiarvoa rajana, tässä tutkielmassa käytettiin kanavan absoluuttisia arvoja. Yksi raja-arvo ei kuitenkaan toiminut kaikilla vuosilla johtuen eroista kuvien ominaisuuksien välillä. Varjojen syvyys vaihteli vuosien välillä, jolloin sama arvo olisi rajannut liikaa yhtenä vuotena ja liian vähän toisena. Käytetyt raja-arvot järjestyksessä vuodesta 2005 alkaen olivat: 20, 30, 40, 30, 30, 25 sekä 30. Lisäksi myös sekä vihreän että sinisen kanavan sävyarvojen tuli olla suurempia kuin nolla, jotta pikseliä hyödynnettiin piirrelaskennassa.



Kuva 8. Vasemmalla kuollut esimerkipuu vuodelta 2013 ja oikealla vuodelta 2017. Yllä näkyvät kaikki latvussegmentin rajaamat pikselit ja alla varjojen rajauksen jälkeen piirteiden laskentaan käytetyt pikselit.

2.8 Kasvillisuusindeksien laskenta

Jokaiselle pikselille laskettiin kolmesta kanavasta erilaisia kasvillisuusindeksejä. Yksi eniten käytetty ja hyväksi havaittu indeksi kuolleiden puiden kartoituksessa on punaisesta ja vihreästä kanavasta laskettu *RGI* (Coops ym. 2006, Meddens ym. 2011, Hart ja Veblen

2015). Muita laskettuja indeksejä olivat muun muassa kanavien normalisoidut eroavaisuusindeksit, esimerkiksi *NDRG* sekä valaistuseroja ja ilmakehän vaikutuksia vähentävä *VARI* (Gitelson ym. 2002, Freeman ym. 2016). *VARI* on kehitetty etenkin RGB-kuville, koska sen laskennassa hyödynnetään kaikkia kolmea kanavaa. Kaikki lasketut indeksit sekä niiden kaavat löytyvät taulukosta 3.

Taulukko 3. Kanavista lasketut kasvillisuusindeksit sekä niiden kaavat.

Indeksi	Kuvaus	Kaava
<i>RGI</i>	Red-Green index	$\frac{Red}{Green}$
<i>RBI</i>	Red-Blue index	$\frac{Red}{Blue}$
<i>GRI</i>	Green-Red index	$\frac{Green}{Red}$
<i>GBI</i>	Green-Blue index	$\frac{Green}{Blue}$
<i>BRI</i>	Blue-Red index	$\frac{Blue}{Red}$
<i>BGI</i>	Blue-Green index	$\frac{Blue}{Green}$
<i>R_G</i>	Punaisen ja vihreän kanavan erotus	$Red - Green$
<i>R_B</i>	Punaisen ja sinisen kanavan erotus	$Red - Blue$
<i>G_B</i>	Vihreän ja sinisen kanavan erotus	$Green - Blue$
<i>NDRG</i>	Normalized Difference Red-Green index	$\frac{(Red - Green)}{(Red + Green)}$
<i>NDRB</i>	Normalized Difference Red-Blue index	$\frac{(Red - Blue)}{(Red + Blue)}$
<i>NDGB</i>	Normalized Difference Green-Blue index	$\frac{(Green - Blue)}{(Green + Blue)}$
<i>VARI</i>	Visible Atmospherically Resistant index	$\frac{(Green - Red)}{(Green + Red - Blue)}$

Indeksien laskennan jälkeen jokaiselle puulle laskettiin kanavien sävyarvoista sekä indekseistä tilastollisia muuttujia. Lasketut tilastolliset muuttujat näkyvät taulukossa 4. Nämä muuttujat olivat yksittäisiä puuta kuvaavia selittäviä muuttujia eli kuvapiirteitä. Kaikkiaan jokaista puuta kuvaamaan saatiin 352 erillistä piirrettä. Punaisen kanavan minimiarvo (*R_min*) jätettiin pois piirrejoukosta, koska piirrelaskentaan valitut pikselit oli

rajattu punaisen kanavan raja-arvolla, ja näin ollen kaikilla puilla yksittäisenä vuonna oli sama R_{min} . Mitään muita piirrevalintoja ei tehty ennen puoliautomaattisia luokitteluja.

Taulukko 4. Jokaiselle kanavalle sekä indekseille lasketut tilastolliset muuttujat.

Piirre	Kuvaus
i_{mean}	Keskiarvo
i_{std}	Keskihajonta
i_{min}	Minimi
i_{max}	Maksimi
i_{range}	Maksimin ja minimin erotus
i_{kur}	Kurtoosi eli jakauman huipukkuus
i_{skew}	Jakauman vinous
$i_{p05}, .. i_{p25}, .. i_{p75}, .. i_{p95}$	Persentiilit välillä 05-95, esimerkkien lisäksi tasakymmenet
$i_{D05}, i_{D25}, i_{D50}, i_{D75}$	Tiheysmuuttujat: 95. persentiili jaettuna 5., 25., 50. sekä 75. persentiilillä

2.9 Puiden puoliautomaattinen luokittelu

Puoliautomaattiset luokittelut toteutettiin R-ohjelmalla käyttäen neljää eri koneoppimisalgoritmia. Käytetyt algoritmit olivat Random Forest (RF), lähimmän naapurin menetelmää hyödyntävä Random Forest (RF+kNN), logistinen regressio (LR, logistic regression) sekä lineaarinen diskriminanttianalyysi (LDA, linear discriminant analysis). Jatkossa käytettyihin koneoppimismenetelmiin viitataan suluissa näkyvillä lyhenteillä.

Random Forest (RF) on regressiopuihin perustuva koneoppimismenetelmä. (Breiman 2001). Algoritmi muodostaa sille annetun aineiston selittävistä muuttujista regressiopuupareja, näistä puupareista muodostuu ”metsä” ja tulos valitaan enemmistöäänestyksellä. Random Forest -algoritmeilla voidaan myös hyödyntää lähimmän naapurin (kNN, k Nearest Neighbor) menetelmiä. Random Forestin onkin todettu olevan hyvä metsäisten tunnusten laskemiseen muihin lähimmän naapurin menetelmiin verrattuna (Vastaranta ym. 2013.) Lähimpien naapurien lukumäärä vaikuttaa saatavaan tulokseen ja sopivan k-arvon valinta onkin tärkeää (Hudak ym. 2008, Vastaranta ym. 2013.) Oletusarvona käytetään

tetyssä menetelmässä k :n arvo oli 5 ja useamman vuoden aineistoilla tehtyjen testien perusteella se myös valittiin lopulliseksi arvoksi. Lähimmän naapurin Random Forest luokittelu toteutettiin ”yaimpute” -paketilla (Crookston ja Finley 2008).

Lineaarinen diskriminanttianalyysi (LDA) on lineaariseen regressioon pohjautuva yksinkertainen luokittelumenetelmä. Havaintojen luokittelu tapahtuu lähimmän luokan keskipisteen perusteella (Hast 2010). Yksinkertaisuutensa vuoksi LDA ei pysty aina havainnoimaan luokkien eroja riittävällä tarkkuudella ja luokittelun tarkkuus heikkenee selittävien muuttujien määrän kasvaessa (Hast 2010).

Logistisessa regressiossa (LR) ennustetaan selitettävän tapahtuman esimerkiksi ennalta määritettyjen luokkien todennäköisyyksiä. Käytettyjen luokkien lukumäärän mukaan vaihtoehtoina ovat kahden luokan binäärinen sekä useamman luokan multinominen logistinen regressio. Logistinen regressio toteutettiin tässä tutkielmassa ”glmnet” -paketin (Friedman ym. 2021) funktioilla. Kyseinen paketti soveltuu hyvin käsittelemään keskenään korreloivia selittäviä muuttujia sekä aineistoja, joissa havaintojen osuus on selittävien muuttujien lukumäärää pienempi (Hastie ym. 2021). Funktio etsii parhaat selittävät muuttujat ristiin validoinnilla, ja osaa pakottaa huonojen selittävien muuttujien tärkeyden pieneksi (Hastie ym. 2021). Puiden terveydentilan luokittelussa käytettiin multinomista logistista regressiota, mutta vuoden 2005 aineistossa oli vain terveitä ja kuolleita puita, joten kyseisenä vuonna käytettiin binääristä regressiota.

Puoliautomaattisessa luokittelussa koneoppimismenetelmä opetetaan tekemään haluttu luokittelu osalla aineistoa ja luokittelun tarkkuutta voidaan arvioida käyttämättömällä luokittelumenetelmälle täysin uudella testiaineistolla. Tässä tutkielmassa aineisto jaettiin koulutus- ja testiaineistoihin suhteessa 70/30. Jaossa huomioitiin eri luokat ja niiden epätasapaino toteuttamalla jako jokaiselle terveydentilaluokalle erikseen. Näin saatiin kaikista terveistä, sairaista sekä kuolleista puista tasan 70 % koulutusaineistoon. Selittävinä piirteinä käytettiin kaikkia laskettuja kuvapiirteitä ja selitettävänä muuttujana puun terveydentilaa.

Luokittelut toteutettiin kaikkina seitsemänä vuotena kaikilla neljällä menetelmällä. Lisäksi jokainen luokittelu sekä aineiston jako satunnaisesti opetus- ja testiaineistoihin tois-

tettiin kymmenen kertaa. Näistä kymmenestä toistosta voitiin laskea keskiarvot. Näin saatiin tuloksiin varmuutta, sillä jokainen luokittelu on satunnainen ja tulokset vaihtelevat hieman jokaisella kerralla aineistojen mukaan. Vastaavissa tutkimuksissa luokittelu on toistettu esimerkiksi viisi kertaa, mutta valittu määrä on aina jossain määrin laskennallisen tehokkuuden ja satunnaisuuden vähentämisen kompromissi (Krzystek ym. 2020).

2.10 Luokittelun tarkkuuden arviointi

Luokittelun tarkkuutta ja onnistumista tarkasteltiin kokonaistarkkuudella, kappa-arvolla (Cohen's kappa, Cohen 1960) sekä konfuusiomatriiseilla. Kokonaistarkkuus kertoo kaikkien oikeiden menneiden luokittelujen osuuden kokonaismäärästä. Kappa-arvolla (Yhtälö 1) voidaan arvioida luokittelun laatua, sillä se ottaa myös sattuman huomioon. Kappa saa arvoja välillä -1–1, tosin negatiiviset arvot ovat harvinaisia (Sirén 2008). Negatiivinen arvo tarkoittaisi, että luokittelu on huonompi kuin sattumalta saatu tulos. Kappa-arvo 1 kertoo täydestä yksimielisyydestä luokittelijoiden kesken, kun taas arvo 0 kertoo tulosten voivan johtua pelkästä sattumasta (McHugh 2012).

$$K = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)} \quad (1),$$

jossa $\Pr(a)$ on suhteellinen osuus, jolloin arvioijat ovat yksimielisiä, $\Pr(e)$ on sattumayksimielisyyden suuruus olettaen, että päätökset ovat riippumattomia.

Kappa-arvo onkin hyvä arvioimaan todellista tarkkuutta, sillä se osaa erottaa sattumalta oikein menneet tulokset (Meddens ym. 2011). Perinteisesti alle 0,6 kappa-arvoja on pidetty korkeintaan keskinkertaisina, välillä 0,6-0,8 huomattavina ja arvon ollessa yli 0,8 lähes täydellisinä (Taulukko 5) (Landis ja Koch 1977).

Taulukko 5. Kappa-arvojen tulkinta (Landis ja Koch 1977)

Kappa-arvo	Yksimielisyyden voimakkuus
<0,00	Huono
0,00–0,20	Heikko
0,21–0,40	Kohtalainen
0,41–0,60	Keskinkertainen
0,61–0,80	Huomattava
0,81–1,00	Lähes täydellinen

Kaikista luokittelutuloksista muodostettiin konfuusiomatriisit (Kuva 9). Konfuusiomatriiseilla voitiin tarkastella luokittelutuloksia ja tarkkuuksia luokkien välillä. Niistä nähtiin myös mihin luokkiin väärät tulokset painoutuivat. Värillä korostetut ruudut näyttävän jokaisen luokan oikeinluokittelutarkkuuden eli sen, kuinka monta puuta tai prosenttia kyseisestä luokasta luokiteltiin oikein. Vuoden 2019 esimerkissä terveistä puista luokiteltiin oikein 162 kappaletta, 98 % kaikista terveistä puista. Vastaavasti valkoisista ruuduista voidaan tarvittaessa tarkastella mihin luokkaan väärin luokitellut sekoittuivat.

		Tarkkuus 96.6% Kappa 0.93 2005 RF+kNN			Tarkkuus 93.8% Kappa 0.9 2019 RF		
		Referenssi			Referenssi		
		Terve	Kuollut		Terve	Sairas	Kuollut
Ennuste	Terve	29 (94%)	0 (0%)		162 (98%)	3 (7%)	9 (6%)
	Kuollut	2 (6%)	27 (100%)		1 (1%)	39 (87%)	4 (2%)
					3 (2%)	3 (7%)	147 (92%)

Kuva 9. Esimerkit puiden terveydentilan luokittelusta muodostetuista konfuusiomatriiseista, vasemmalla vuodelta 2005 ja oikealla 2019. Värilliset ruudut näyttävät jokaisen luokan oikeinluokittelutarkkuuden, ylhäällä näkyy luokittelun kokonaistarkkuus, kappa-arvo sekä käytetty luokittelumenetelmä.

3. TULOKSET

3.1 Kuolleisuuden kehitys keskuspuiston alueella tutkimusjakson aikana

Taulukossa 6 näkyvät eri vuosien sairaiden ja kuolleiden puiden lukumäärät. Vuosi 2019 jätettiin pois tuloksista, sillä määrät eivät olleet vertailukelpoisia muiden vuosien kanssa. Yhteensä kriteerien mukaisia sairaita puita löytyi 473 ja kuolleita puita 6 227. Kuolleiden puiden määrä ei kuitenkaan kerro niiden todellista määrää, sillä osa puista oli näkyvissä useamman vuoden kuvassa. Edeltävien vuosien keskilämpötilojen keskiarvo korreloi vahvasti kuolleisuuden kanssa korrelaatiokertoimen ollessa noin 0,73 ($p=0,09$). Sademäärillä vastaavaa riippuvuutta ei ollut havaittavissa korrelaatiokertoimen ollessa -0,04 ($p=0,94$).

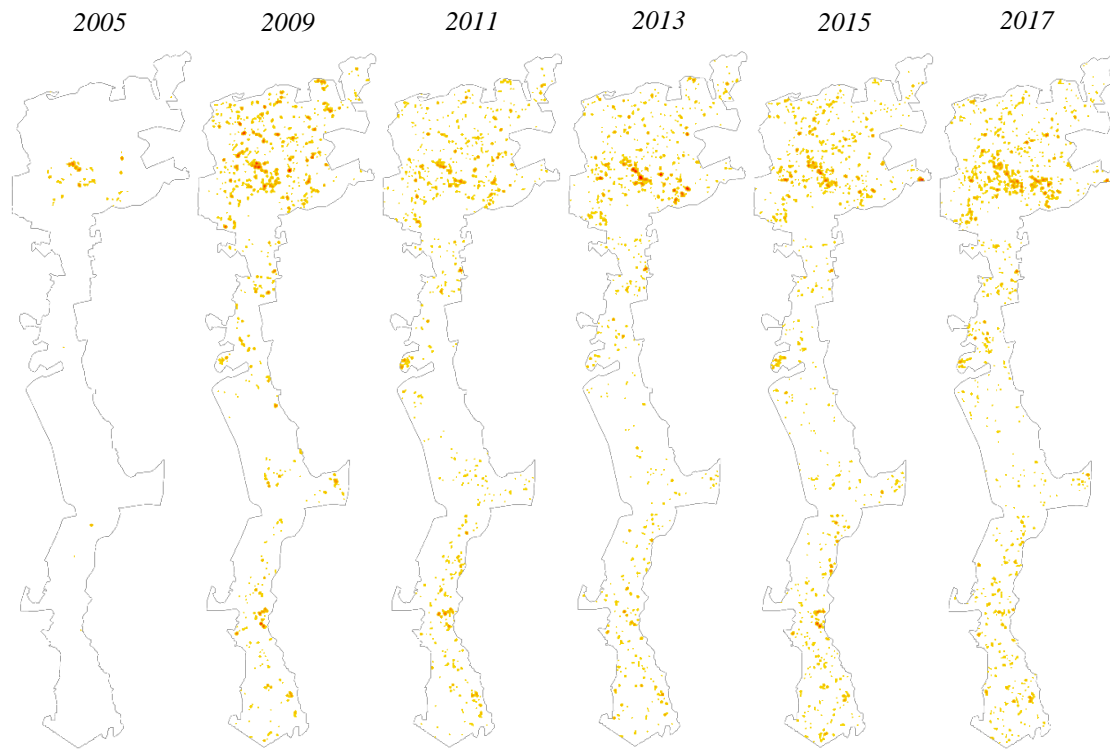
Vuonna 2005 kuolleita puita löytyi erittäin vähän (113) ja vuonna 2009 taas toiseksi eniten (1 339). Kuvien välissä oli muista vuosista poiketen neljän vuoden väli, mikä pitää huomioida tarkastellessa 2009 määriä. Vuoden 2011 jälkeen kuolleiden puiden määrissä oli havaittavissa nousevaa trendiä ja huippu saavutettiin vuonna 2017. Sairaita puita löytyi vuosien välillä hyvin vaihtelevasti, pienimmillään 0 ja suurimmillaan 195. Kaikki sairaat puut olivat havupuita, sillä heikkenemisen tarkastelu värioireiden perusteella perustuu neulasten väreihin kuolemisprosessin eri vaiheissa.

Taulukko 6. Sairaiden ja kuolleiden puiden lukumäärät Helsingin keskuspuiston alueella.

Vuosi	Sairaat	Kuolleet	Kuolleista lehtipuita
2005	0	113	0
2009	58	1339	6
2011	195	1035	12
2013	47	1128	21
2015	155	1212	83
2017	18	1400	74
Yhteensä	473	6 227	196

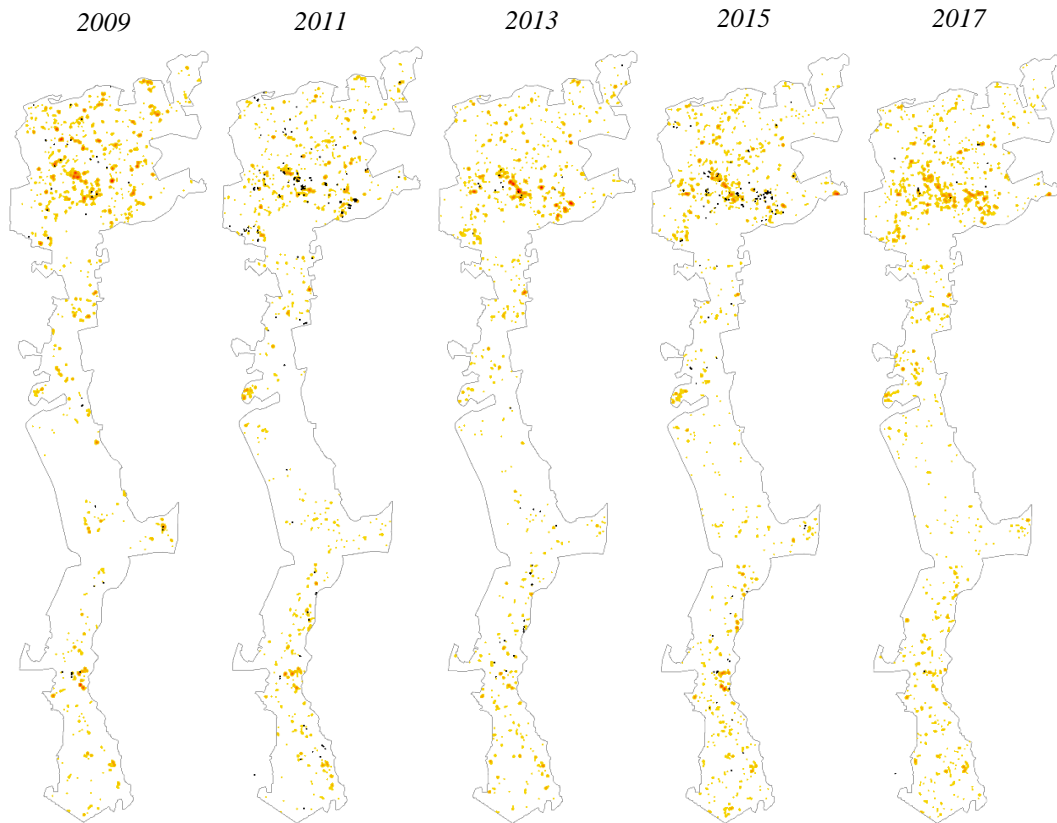
Lämpökartta-aikasarjassa (Kuva 10) näkyy puiden kuolleisuuden määrällinen ja alueellinen vaihtelu 13 vuoden aikajänteellä. Vuosien välillä voidaan nähdä selkeitä eroja, ja kuten taulukosta 5 nähdään vuonna 2009 oli toiseksi eniten kuolleita puita määrällisesti.

Tuolloin kuolleet puut eivät kuitenkaan olleet levinneet yhtä laajalle alueelle kuin myöhemmin. Ne olivat enemmänkin keskittyneitä pienille alueille, jotka näkyvät kuvassa 10 hyvin tummina alueina. Vuoden 2009 jälkeen muutaman seuraavan vuoden määrät nousivat pikkuhiljaa ja kuolleita puita alkoi olla laajemmalla alueella. Laajimmillaan sekä alueellisesti että määrällisesti puiden kuolleisuus oli vuonna 2017, jolloin kuvassa voidaan nähdä kuolleita puita uusilla alueilla.



Kuva 10. Kuolleiden puiden sijainneista luotu lämpökartta-aikasarja Helsingin keskuspuiston alueelta. Yksittäiset kuolleet puut näkyvät oransseina pisteinä kartalla, useamman kuolleen puun ryhmät sitä tummemman punaisena mitä enemmän kuolleita puita oli.

Kuvassa 11 näkyy kuolleisuuden lisäksi sairaiden puiden sijainnit. Vuoden 2005 aineistossa ei ollut sairaita puita, joten se jätettiin aikasarjasta pois. Vuoden 2009 sairaita puita oli vain yksitellen, mutta 2011 niitä alkoi olla enemmän samoissa ryppäissä. Aikasarjasta huomataan etenkin vuoden 2015 sairaiden puiden painottuvan samalla alueella. Aiempina vuosina kyseisellä alueella ei ole ollut isoja määriä kuolleita puita. Kokonaisuutena erot vuosien välillä ovat suuria sairaiden puiden määrien alueellisissa muutoksissa. Kaikki aikasarjojen kuvat löytyvät suurempina versioina liitteistä 1-6.

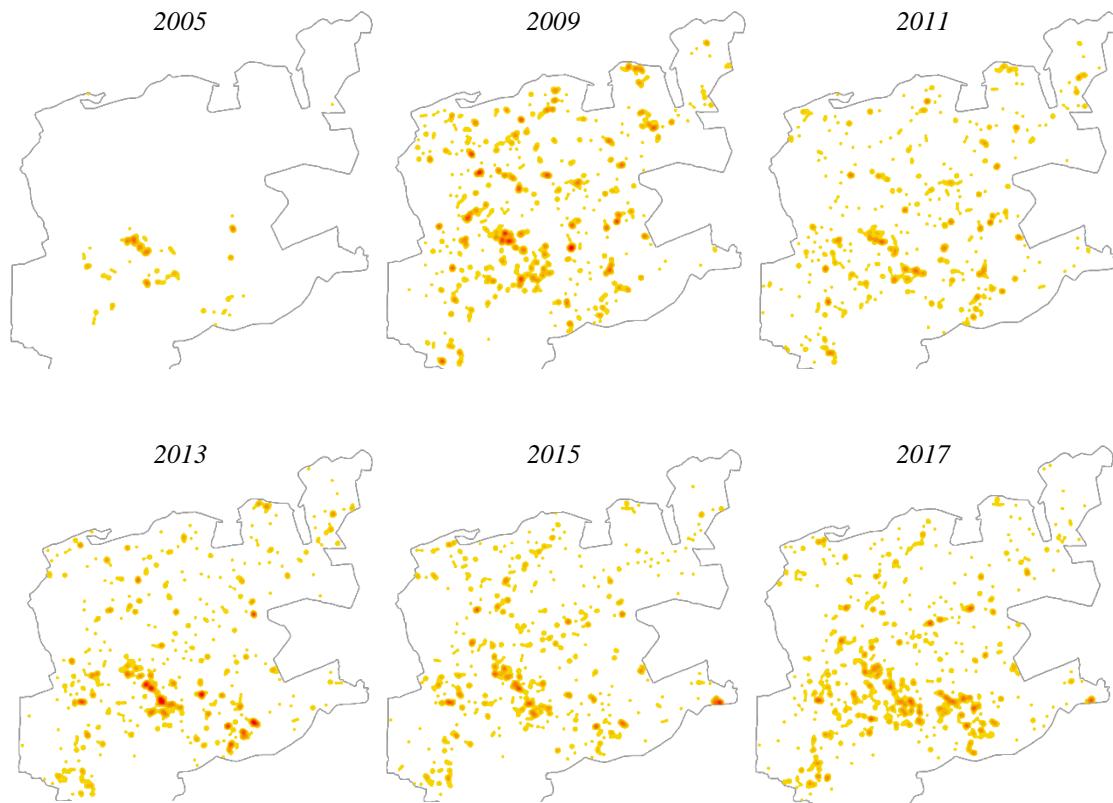


Kuva 11. Kuolleiden puiden sijainneista luotu lämpökartta-aikasarja Helsingin keskuspuiston alueelta, johon on lisätty sairaiden puiden sijainnit mustilla pisteillä. Yksittäiset kuolleet puut näkyvät oransseina pisteinä kartalla, useamman kuolleen puun ryhmät sitä tummemman punaisena mitä enemmän kuolleita puita oli.

3.2 Kuolleisuuden kehitys keskuspuiston pohjoisosassa

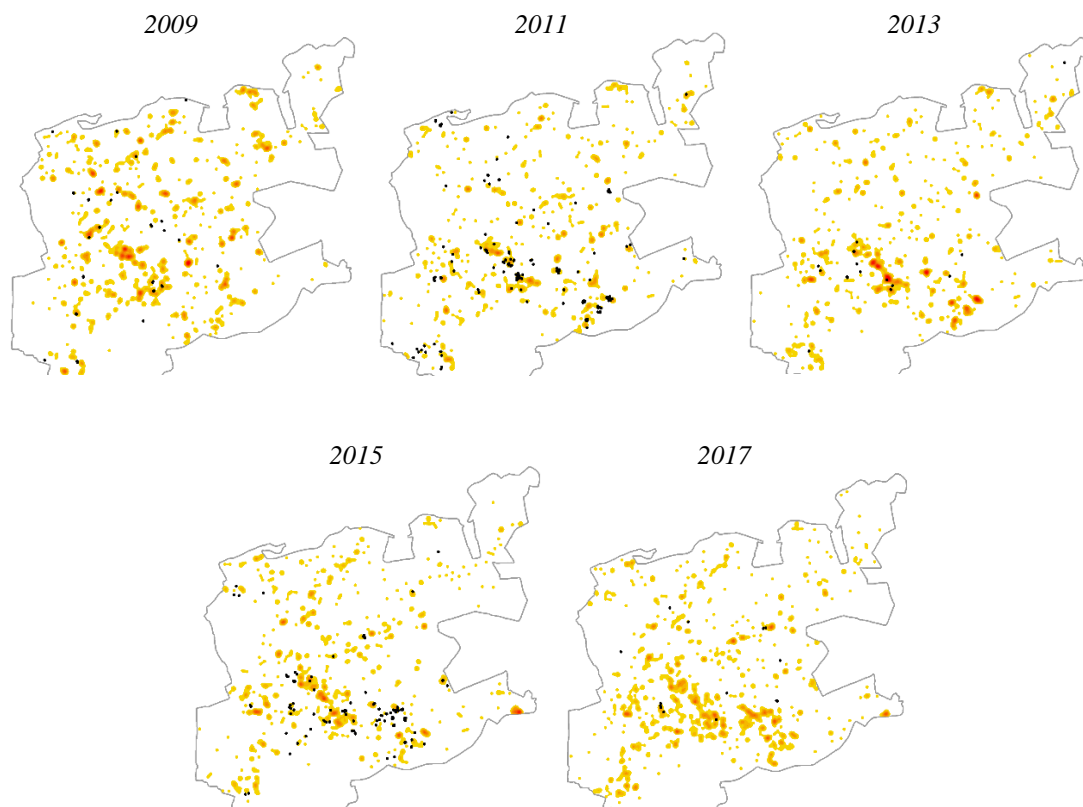
Suurimmat muutokset puiden kuolleisuudessa sekä alueellisesti että määrällisesti ovat tapahtuneet keskuspuiston pohjoisosassa (Kuva 12). Kuolleita puita löytyi vuonna 2005 vain hyvin pieneltä alueelta, mutta jo vuonna 2009 niitä löytyi käytännössä koko näkyvissä olevalta alueelta. Tuolloin oli myös hyvin selkeästi havaittavia tummia ryppäitä, joissa kuolleita puita oli runsaasti. Vuoteen 2011 mennessä tilanne rauhoittui selkeästi ja etenkin pahimmat alueet näkyvät aiempaa vaaleampina. Kuolleiden puiden lukumäärä lähti hienoiseen nousuun vuoden 2011 jälkeen, ja kuolleisuus kasvoi etenkin aiemmin muodostuneissa kuolleiden puiden ryppäissä. Nämä alueet näkyvät hyvin tummina kohdina vuonna 2013. Vuosina 2015 ja 2017 kuolleiden puiden lukumäärä oli vuotta 2013 suurempi, mutta kyseisinä vuosina ei ollut havaittavissa yhtä tummia ryppäitä. Kuolleet

puut olivat jakautuneet laajemmalle alueelle ja vuonna 2017 kuolleita puuryhmiä löytyi jo hyvin laajasti ympäri kuvassa 12 näkyvissä olevaa aluetta.



Kuva 12. Kuolleiden puiden sijainneista luotu lämpökartta-aikasarja Helsingin keskuspuiston pohjoisosasta. Yksittäiset kuolleet puut näkyvät oransseina pisteinä kartalla, useamman kuolleen puun ryhmät sitä tummemman punaisena mitä enemmän kuolleita puita oli.

Myös sairaiden puiden määrissä suurimmat muutokset olivat näkyvissä puiston pohjoisosassa (Kuva 13). Erot vuosien välillä olivat suuret ja vuosina 2009, 2013 sekä 2017 sairaiden puiden määrät olivat hyvin pieniä ja puut olivat hajallaan toisistaan. Vuonna 2011 löytyi yksittäisten sairaiden puiden lisäksi hyvin tiiviitä sairaiden puiden ryppäitä. Lisäksi nämä puuryhmät olivat alueilla, joilla oli jo ennestään useampia kuolleita puita. Vuoden 2015 sairaista puista iso osa sijaitsi samassa rykelmässä, ja huomionarvoista sijainnissa oli myös se, että aiemmin alueen kuolleiden puiden määrä oli pysynyt maltillisena. Vuonna 2017 kyseiseltä alueelta löytyi selvästi enemmän kuolleita puita.



Kuva 13. Kuolleiden puiden sijainneista luotu lämpökartta-aikasarja Helsingin keskuspuiston pohjoisosasta, johon on lisätty sairaiden puiden sijainnit mustilla pisteillä. Yksittäiset kuolleet puut näkyvät oransseina pisteinä kartalla, useamman kuolleen puun ryhmät sitä tummemman punaisena mitä enemmän kuolleita puita oli.

3.3 Luokittelutarkkuus terveydentilan mukaan

Keskimääräinen luokittelutarkkuus kolmen luokan luokittelussa (terve, sairas, kuollut) kaikilla menetelmillä kaikkina vuosina oli noin 94 % (Taulukko 7) ja kappa-arvo noin 0,88. Vain vuoden 2015 kokonaistarkkuuden keskiarvo (89,4 %) jäi alle 90 % tarkkuuden. Paras luokittelutarkkuus saavutettiin niukasti vuonna 2013 (95,6 %) ja toiseksi paras vuonna 2017 (95,56 %). Keskimäärin voidaan sanoa lähes kaikkien vuosien luokittelutarkkuuksien keskiarvojen olevan hyvin lähellä toisiaan.

Luokittelumenetelmistä parhaiten pärjasi LR, sillä saavutettiin paras luokittelutarkkuus kaikkina muina vuosina paitsi 2005. Paras yksittäinen luokittelutarkkuus oli 96,9 % (RF+kNN), tosin samana vuonna (2005) oli muitakin hyviä tuloksia. Vuoden 2005 aineisto oli kooltaan muita vuosia selvästi pienempi. Luokittelutarkkuuksissa kahtiajako oli

selvä, LR sai parhaimpia tuloksia ja jokaisena vuonna heikoimmin suoriutuivat RF+kNN tai LDA. Etenkin LDA:lla muutamien tulokset olivat selkeästi muiden menetelmien tuloksia huonompia, esimerkkeinä vuodet 2005 sekä 2019. Toisaalta vuonna 2017 kolme menetelmää oli hyvin lähellä toisiaan.

Taulukko 7. Puiden terveydentilaluokittelun kokonaistarkkuudet vuosittain ja menetelmäkohtaisesti.

Vuosi	Keskiarvo-tarkkuus, %	
2005	94,7	RF+kNN (96,9) LR (96,7) RF (96,6) LDA (88,4)
2009	93,8	LR (94,5) LDA (94,2) RF (93,4) RF+kNN (93,0)
2011	92,8	LR (93,3) LDA (92,8) RF (92,7) RF+kNN (92,5)
2013	95,6	LR (95,9) RF (95,8) RF+kNN (95,6) LDA (95,1)
2015	89,4	LR (90,1) LDA (89,2) RF (89,1) RF+kNN (89,0)
2017	95,56	LR (96,0) RF+kNN (95,44) RF (95,42) LDA (95,41)
2019	93,7	LR (94,4) RF+kNN (93,9) RF (93,8) LDA (92,6)
Vuosien keskiarvo	93,6	

Kappa-arvoja tarkastellessa kaikkien vuosien keskiarvo oli noin 0,88 (Taulukko 8). Paras kappa-arvon keskiarvo 0,914 oli vuonna 2013, samana vuonna oli myös paras kokonaistarkkuus. Heikoin tulos sekä luokittelutarkkuuksissa että kappa-arvoissa saavutettiin vuoden 2015 aineistolla. Eri luokittelumenetelmien järjestys säilyi samankaltaisena kuin taulukossa 7, parhaan arvon melkein kaikkina vuosina saavutti LR. Heikoiten suoriutuivat RF+kNN sekä LDA. Lähes kaikkina vuosina etenkin heikommin suoriutuneiden menetelmien tulokset olivat lähellä toisiaan. Poikkeuksena LDA:n tulos vuoden 2005 aineistolla, jossa tulos poikkesi merkittävästi muiden menetelmien tuloksista.

Taulukko 8. Puiden terveydentilaluokittelun kappa-arvot vuosittain ja menetelmäkohtaisesti.

Vuosi	Keskiarvo-kappa	
2005	0,89	RF+kNN (0,938) LR (0,934) RF (0,931) LDA (0,767)
2009	0,88	LR (0,891) LDA (0,885) RF (0,869) RF+kNN (0,861)
2011	0,87	LR (0,881) LDA (0,872) RF (0,868) RF+kNN (0,865)
2013	0,914	LR (0,919) RF (0,917) RF+kNN (0,913) LDA (0,906)
2015	0,80	LR (0,818) LDA (0,803) RF (0,799) RF+kNN (0,798)
2017	0,912	LR (0,921) LDA (0,91) RF+kNN (0,9095) RF (0,909)
2019	0,89	LR (0,908) RF+kNN (0,898) RF (0,896) LDA (0,876)
Vuosien keskiarvo	0,88	

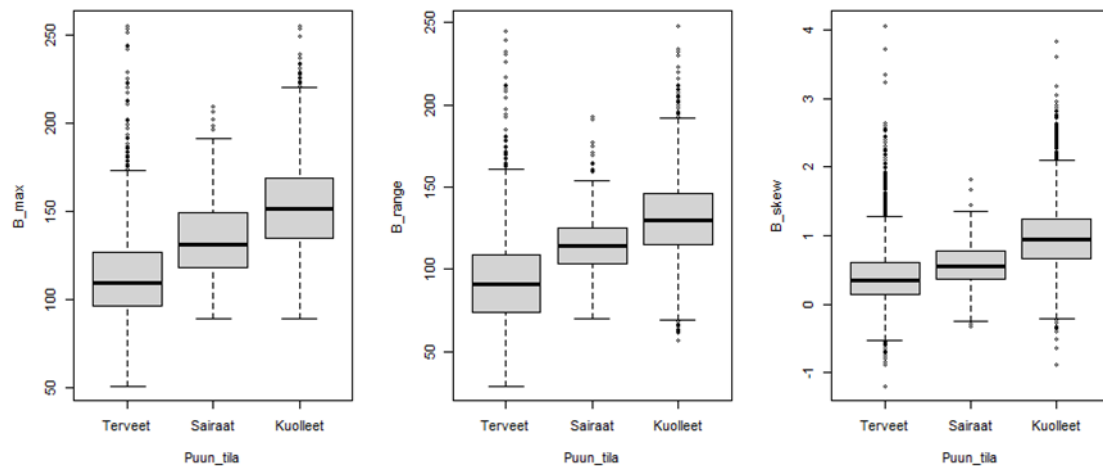
Taulukossa 9 näkyvät kaikkien terveydentilaluokittelujen konfuusiomatriisien oikeinluokittelutarkkuudet. Yksittäisestä luokittelutuloksesta on näkyvissä vain oikein menneiden osuus (Kuvan 9 konfuusiomatriisin värilliset ruudut). Esimerkiksi vuoden 2005 ensimmäisestä luokittelusta (RF+kNN) nähdään, että terveistä puista luokiteltiin oikein 96 %, sairaista 0 % ja kuolleista 99 %. Oikeassa reunassa nähdään lisäksi luokittelun arviointiin käytetyn testiaineiston jakauma, eli missä suhteessa terveitä ja kuolleita puita on keskenään sekä verrattuna sairaisiin puihin.

Tuloksista nähdään sairaiden puiden tulosten vaihtelu. Paras tulos oli 94 % huonoimman ollessa 0 %. Terveillä puilla kaikki tulokset olivat yli 90 %, kuolleilla ei päästy aivan samaan tulokseen. Muutamassa tapauksessa oikeinluokitteluprosentti oli alle 90 %. Lähes kaikissa luokitteluissa terveiden puiden oikeinluokitteluprosentti oli kuolleita parempi. Testiaineiston sisältämä puiden lukumäärä vaihteli 58–810 välillä. Terveiden ja kuolleiden puiden keskinäiset suhteet olivat lähellä toisiaan. Sairaita puita oli vähemmän, 2005 ei yhtään ja 2019 eniten (45).

Taulukko 9. Kaikkien terveydentilaluokkien oikeinluokittelutarkkuudet vuosittain ja menetelmäkohtaisesti sekä testiaineiston jakauma luokkien kesken.

	RF+kNN	RF	LR	LDA	
Vuosi	Terve/sairas/kuollut %	Terve/sairas/kuollut %	Terve/sairas/kuollut %	Terve/sairas/kuollut %	Terveet+sairaat+kuolleet=yht kpl
2005	96/0/99	96/0/98	97/0/96	92/0/84	31+0+27 = 58
2009	96/2/92	96/2/93	97/19/94	98/31/92	389+10+309 = 708
2011	95/65/94	95/61/95	95/74/94	95/75/94	302+42+259 = 603
2013	98/8/97	97/6/97	98/6/97	97/32/95	327+11+288 = 626
2015	92/61/89	92/56/90	94/55/90	94/58/88	361+40+337 = 738
2017	95/0/97	96/0/97	97/4/96	97/18/95	415+5+390 = 810
2019	96/94/92	96/92/92	97/90/93	97/79/92	166+45+160 = 371

Luokitteluissa käytetyistä piirteistä parhaat olivat samoja sekä vuosien välillä että myös vuosien sisällä. Kaikilla menetelmillä sinisen kanavan piirteet nousivat käytetyimmäksi. Kolme parasta piirrettä olivat sinisen kanavan maksimiarvo (B_{max}), sinisen kanavan maksimin ja minimin erotus (B_{range}) sekä sinisen kanavan arvojen vinous (B_{skew}). Nämä olivat ylivoimaisesti käytetyimpiä ja ne toistuivat kaikilla menetelmillä kaikkina vuosina. Kuvassa 14 näkyy näiden kolmen piirteen boxplot -kaaviot. Kaaviot laskettiin koko luokitteluaineistosta ($n = 13\,047$) ja niistä nähdään silti selvästi erot luokkien välillä.



Kuva 14. Puiden terveydentilaluokittelun kolmen tärkeimmän piirteen boxplot -kaaviot. Kaaviossa jana kuvaa pienimpiä ja suurimpia arvoja, laatikon ala- ja yläreuna 25. ja 75. persenttiilejä sekä laatikon keskiviiva mediaania. Pisteet ovat poikkeaviksi tulkittuja ääriarvoja. Kaaviot laskettiin koko luokitteluaineistosta, puiden lukumäärä oli 13 047.

4. TULOSTEN TARKASTELU

4.1 Kuolleisuuden kehityksen tarkastelu

Tutkielman ensimmäisen puoliskon tarkoituksena oli selvittää, millaisia muutoksia puiden kuolleisuudessa on tapahtunut tutkimusalueella tutkimusjakson aikana (1. tutkimuskysymys). Lisäksi selvitettiin, onko kuolleisuus lisääntynyt tutkimusjakson aikana (2. tutkimuskysymys). Seuranta toteutettiin ilmakuva-aikasarjalla ja tavoitteena oli myös saada selville, miten aikasarjaa voidaan hyödyntää puiden kuolleisuuden seurannassa.

Merkittävimmät muutokset puiden kuolleisuudessa keskuspuiston alueella tapahtuivat vuosien 2005 ja 2009 välillä. Kuolleisuus kasvoi tuona ajanjaksona 2005 vuoden lähtötasoon nähden lähes räjähdysmäisesti. Vuoden 2009 jälkeen kuolleiden puiden määrät pienivät, ja seuraavina vuosina oli havaittavissa nouseva trendi, joka huipentui vuonna 2017. Kokonaisuutena kuolleisuuden havaittiin lisääntyneen, mutta kasvu ei ollut jatkuvaa, koska vuosien välillä oli vaihtelua. Valtaosa kuolleista puista oli havupuita, mikä on loogista, koska alueella on ongelmia kaarnakuoriaisten kanssa. Lisäksi kuolleita lehtipuita on mahdollisesti hankalampi tunnistaa ilmakuvilta, ja ne eivät välttämättä pysy pysyissä yhtä pitkään.

Vuodesta 2005 vuoteen 2009 tapahtuneeseen huomattavaan muutokseen oli todennäköisesti useampiakin syitä, mutta yksi niitä saattoi olla vuoden 2006 kesä. Kyseinen kesä oli poikkeuksellisen kuiva ja Helsingissä kesä oli vuonna 1845 alkaneen mittaushistorian kuivin (Ilmatieteen laitos 2006). Kuivuus on voinut hyvin olla kasvaneen kuolleisuuden tärkeimpiä tekijöitä, jonka seurauksena hyönteistuhot ovat lisääntyneet kaarnakuoriaisten hyötyessä heikentyneistä ja kuolleista puista.

Vuoden 2009 jälkeen kuolleisuus ei kuitenkaan kasvanut vaan pieneni ja vuoteen 2013 mennessäkin kasvu oli maltillista. Kesät 2010 ja 2011 olivat lämpimiä ja niiden seurauksena kirjanpainajalla oli kaksi sukupolvea Etelä-Suomessa (Kankaanhuhta ym. 2014), mutta vuosien 2011 ja 2013 kuolleiden puiden määrissä tämä ei ollut juurikaan havaitta-

vissa. On mahdollista, että alueen ennestään korkea kuolleiden puiden määrä on hyödyttänyt kirjanpajan luontaisia vihollisia, jotka ovat kyenneet osittain hillitsemään kirjanpajajakannan kasvua.

Vuoden 2019 kuolleiden puiden määrät eivät olleet suoraan verrattavissa aiempiin vuosiin johtuen kyseisen vuoden ilmakuvan ominaisuuksista. Osa havaituista kuolleista puista näkyi kuvassa vain osittain esimerkiksi varjojen takia. Aiemmin mainittujen kriteerien mukaisesti jokaisen vuoden kuvalta etsittiin riittävän selkeästi erottuvia kuolleita puita, ja tavoitteena ei ollut kerätä aineistoon jokaista kuollutta puuta. Näin toimien olisi saatu kattavampi aineisto kerättyä, mutta aineiston jatkoanalysointi puoliautomaattisen luokittelun kannalta olisi onnistunut todennäköisesti heikommin. Vuoden 2019 kuolleiden puiden todellisen lukumäärän voitiin kuitenkin todeta olleen vuoden 2017 määrää pienempi. Näin ollen kuolleisuudessa aiemmin havaittu kasvava trendi olisi ainakin hidastunut vuoden 2019 perusteella.

Vertaamalla säähavaintoja kuolleiden puiden lukumääriin voitiin arvioida ilmasto-olosuhteiden vaikutusta puiden kuolleisuuteen. Lasketuilla keskilämpötilojen ja sademäärien keskiarvoilla selvitettiin säähavaintojen ja kuolleiden puiden määrän keskinäistä korrelaatiota. Tuloksista huomattiin keskilämpötilojen korreloivan vahvasti kuolleisuuden kanssa ($r=0,73$), mutta sademäärillä korrelaatiota ei ollut havaittavissa ($r=-0,04$). Tämän perusteella onkin todennäköistä, että keskimäärin vuoden keskilämpötila vaikuttaa vuotuista sademäärää enemmän kuolleiden puiden määrään. Lisäksi nähdään, että korkeammat keskilämpötilat voivat lisätä kuolleisuutta.

Talven 2019–2020 keskilämpötila oli koko Suomessa noin $5\text{ }^{\circ}\text{C}$ korkeampi verrattuna vertailujaksoon vuosilta 1981–2010 (Ilmatieteen laitos 2020). Lisäksi Helsingissä talvi oli ennätyksellinen, sillä keskilämpötila oli talvesta 1829–1830 alkaneen mittaushistorian lämpimin (Ilmatieteen laitos 2020). Pääosin poikkeuksellisen lämpimän talven seurauksena myös koko vuoden lämpötila oli poikkeuksellisen korkea ollen noin $2,8\text{ }^{\circ}\text{C}$ keskiarvoa korkeampi (Ilmatieteen laitos 2021). Tämän tutkielman tutkimusjakson aikana yhdenkään vuoden keskilämpötila ei ollut yhtä korkea kuin vuonna 2020. Nähdyn keskilämpötilojen ja kuolleisuuden riippuvuuden perusteella voidaan pitää todennäköisenä,

että vuoden 2019 jälkeen kuolleisuuden trendi kääntyy jälleen nousevaksi. Tämän seurauksena kuolleiden puiden määrä kasvaisi vuonna 2021, mutta määriä on mahdoton arvioida.

Yhden vuoden ilmakuvalta kuollut puu voidaan tunnistaa, mutta on mahdotonta sanoa mikä sen tappoi tai milloin se kuoli. Kuollut puu voi pysyä pystyssä pitkäänkin, tämänkin tutkielman kuvista löytyi kuolleita puita, jotka näkyivät kuudella eri kuvalla. Näin ollen ne olivat olleet pystyssä ainakin 10 vuotta. Ilmakuva-aikasarjat mahdollistavatkin puiden kuolemisajankohdan määrittämisen. Esimerkiksi puun kuolinvuosi on erittäin hyödyllinen tieto, jolla voidaan tarkastella kuolleiden puiden määrän kehitystä sekä alueellista painopistettä. Aikasarjoilla voidaan muutenkin seurata yksittäisten puiden terveydentilaa, vaikka ne eivät kuolisikaan. Seurannalla voidaan ymmärtää paremmin miksi heikentyneet puut eivät aina kuole, vaan palautuvat terveiksi.

Manuaalisella tarkastelulla aikaisempien vuosien kuvista voidaan luotettavasti päätellä, milloin puu on kuollut. Tätä tietoa ei kuitenkaan saatu riittävällä tarkkuudella kerättyä tutkielman aineistosta. Uusia kuolleita puita ei onnistuttu erottamaan automaattisesti viereisistä vanhemmista kuolleista puista. Aikasarjoihin perustuva yksittäisten puiden seuranta luokien tiettyjä vaatimuksia käytetyille kaukokartoitusaineistoille. Ensinnäkin spatiaalisen resoluution tulisi olla riittävä, jotta puita voidaan tarkastella puutasolla. Toinen tärkeä asia on aineistojen laatu, jotta puiden sijainnit eivät vaihtelee aineistojen välillä. Puiden tulisi näkyä kohtisuoraan ylöspäin ja georeferoinnin olla tarkkaa, jotta kahden ajankohdan vertailussa vertailtavien kohteiden sijainnit olisivat mahdollisimman samoja.

Tämän tutkielman aineistossa oli myös sairaita puita, joita voitiin käyttää paikkaamaan tiedonpuutetta puiden kuolemisajankohdista. Ne eivät olleet vielä kuolleita, mutta olivat kuolemassa ja niitä tarkastelemalla voidaan nähdä, missä kuolleiden puiden määrä on kasvussa. Näin saadaan hyödyllistä informaatiota kuolleisuuden muutoksista, ja siihen voidaan jopa varautua. Vuoden 2015 tuloksissa nähtiin selvästi sairaiden puiden painotuneen samalle alueelle ja vuonna 2017 alueella oli kuolleisuus selvästi kasvussa. Sairaiden puiden tarkastelu ei kuitenkaan aina kerro absoluuttista totuutta puiden kuolleisuuden kehityksestä ja kuten tuloksista nähtiin, muutamana vuonna sairaiden puiden lukumäärä

oli hyvin pieni eikä kyseisistä vuosista voitu tehdä juurikaan johtopäätöksiä. Myös kaukokartoitusaineistojen hankinta-ajankohta voi vaikuttaa sairaiden puiden havaitsemiseen, jolloin yhden ajankohdan aineistoon ei voida luottaa täysin varmasti.

4.2 Puiden terveydentilan luokittelutarkkuuden tarkastelu

Tutkimuksen jälkimmäisen puoliskon tarkoituksena oli selvittää, voidaanko yksittäisiä puita luokitella luotettavasti puoliautomaattisilla koneoppimismenetelmillä terveydentilan mukaan (3. tutkimuskysymys). Lisäksi selvitettiin käytettyjen koneoppimismenetelmien välisiä eroja (4. tutkimuskysymys). Puiden terveydentilan luokittelut toteutettiin seitsemänä eri vuonna ja neljällä eri koneoppimismenetelmällä käyttäen vain ortoilmakuvista laskettuja kuvapiirteitä.

Yksittäisten puiden terveydentila saatiin luokiteltua 93,6 % tarkkuudella (Taulukko 7) kappa-arvon ollessa 0,88 (Taulukko 8). Molemmat arvot ovat seitsemän vuoden keskiarvoja, ja niitä voidaan pitää kohtalaisen hyvinä. Käytettyjen koneoppimismenetelmien paremuusjärjestys säilyi pääosin samankaltaisena sekä oikeinluokittelutarkkuuksien että kappa-arvojen osalta.

Aiemmat samankaltaiset tutkimukset voidaan jakaa kahteen ryhmään sen perusteella, miten luokittelu on toteutettu. Osassa se on tehty samalla tavalla, kuin tässä tutkielmassa luokitellen vain tunnistettujen yksittäisten puiden pikselit ja osassa on luokiteltu koko kuva eli kaikki sen pikselit. Yksittäisten puiden terveydentilan luokittelussa tarkkuudet ovat olleet tasaisia, Freeman ym. (2016) saivat kahdella luokalla (elävä, kuollut) noin 92 % tarkkuuden, Sylvain ym. (2019) pääsivät myös kahdella luokalla noin 94 % tarkkuuteen ja Krzystek ym. (2020) saavuttivat kuolleiden puiden luokittelussa noin 92 % tarkkuuden. Kaikissa kolmessa tutkimuksessa aineistoina käytettiin ortoilmakuvia, mutta Krzystek ym. (2020) hyödynsivät myös laserkeilausaineistoja.

Vastaavasti myös koko kuvan luokittelussa tarkkuudet ovat olleet hyvällä tasolla, Meddens ym. (2011) 90 %, Hart ja Veblen (2015) noin 90 % ja Bright ym. (2020) parhaimmillaan 95 %. Kolmessa viimeksi mainitussa tutkimuksessa luokittelu tehtiin useamman kuin kahden luokan kesken käyttäen pääosin satelliittikuvia. Elävien ja kuolleiden puiden lisäksi luokkia oli muun muassa värioireisille sairaille puille, varjoille sekä puuttomille

alueille. Tässä tutkielmassa saadut tulokset puiden terveydentilan luokittelussa ovat samalla tasolla aiempien tutkimuksien kanssa vaihdellen 89–96 % välillä. Tulokset vertautuvat parhaiten kolmeen ensiksi mainittuun tutkimukseen niiden samankaltaisten menetelmien takia. Huomionarvoista kuitenkin on näiden kolmen tutkimuksen käyttäneen luokitteluissa vain kahta luokkaa (elävä, kuollut), kun tässä tutkielmassa luokittelu toteutettiin kolmella luokalla (terve, sairas, kuollut).

Käytettyä aineistoa voidaan pitää laajana muihin vastaaviin tutkimuksiin verratessa. Useimmissa tapauksissa käytetty aineisto on sisältänyt satoja tai tuhansia puita. Aineiston keruutapa on myös vaihdellut visuaalisesta kuvatulkinna perinteiseen maastoaineistoon. Kuvatulkintana kerätyt aineistot ovat myös vaihdelleet sadoista puista muutamiin tuhansiin (Hart ja Veblen 2015, Freeman ym. 2016). Krzystek ym. (2020) käyttivät tutkimuksessaan noin 10 000 kuvatulkinntapua aineistoa. Omassa luokassaan on Sylvainin ym. (2019) keräämä erittäin laaja 292 000 puun aineisto.

Parhaat tulokset saatiin vuosien 2013 ja 2017 aineistoilla. Luokittelutarkkuudet olivat hyvin lähellä toisiaan (95,6 % ja 95,56 %) ja saavutettuja kappa-arvoja voidaan pitää erinomaisina (0,914 ja 0,912). Molempien vuosien hyviin tuloksiin vaikutti sairaiden puiden vähäinen määrä (Taulukko 9). Sairaiden puiden oikeinluokittelutarkkuus oli heikko, mutta niiden vähäisen määrän takia se ei vaikuttanut kokonaistarkkuuteen. Vuoden 2015 tulokset olivat heikoimpia molemmilla mittareilla (89,4 % ja 0,8), johtuen pääosin muita vuosia heikommasta kuolleiden puiden luokittelusta (Taulukko 9).

Ortoilmakuvien spatiaalisella resoluutiolla ei näyttänyt olevan juurikaan vaikutusta luokittelutuloksiin. Kolme parasta kokonaistarkkuutta oli vuosina 2013, 2017 sekä 2005, vastaavasti näiden vuosien spatiaaliset resoluutiot olivat 20 cm, 8 cm ja 20 cm. Aiemmissä tutkimuksissa on myös huomattu, että paremman resoluution ortoilmakuvat eivät välttämättä paranna luokittelussa saatavia tuloksia (Meddens ym. 2011). Meddens ym. (2011) vertasivat eri spatiaalisia resoluutioita luokitellessaan ortoilmakuvan useampaan luokkaan puiden terveydentilan mukaan. Tutkimuksessa alkuperäisenä aineistona käytettiin ortoilmakuvaa 30 senttimetrin spatiaalisella resoluutiolla. Kuva aggregoitiin kolmeen eri resoluutioon (1,2 m, 2,4 m sekä 4,2 m) ja saavutettuja luokittelutarkkuuksia verrattiin keskenään.

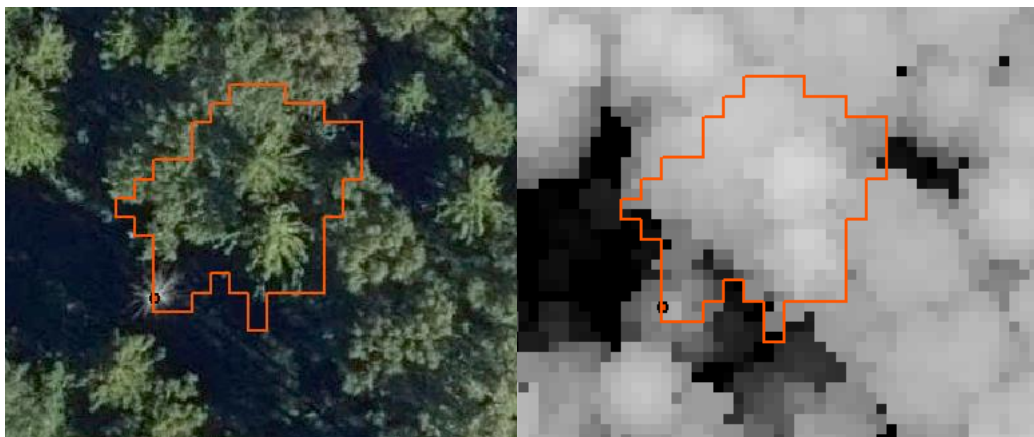
Tuloksista havaittiin parhaan tarkkuuden saavutetun 2,4 metrin spatiaalisella resoluutiolla, ja tuloksista voitiin päätellä luokittelun onnistuvan parhaiten pikselikoon ollessa mahdollisimman samansuuruinen kohteen kanssa. Tutkimuksessa kohteet olivat puiden latvoja, joiden pinta-ala vastasi parhaiten 2,4 metrin pikselikokoa, ja näin puun latvuksen sisällä tapahtuvat sävyarvojen muutokset eivät vaikuttaneet paljoakaan luokittelutarkkuuteen. Pikselikoon ollessa puun latvusta pienempi osuu latvuksen sisälle useita pikseleitä, jolloin yksittäisen puun latvuksessa voi olla sekä aurinkoisia että varjoisia kohtia. Nämä sävyarvojen erot latvuksen sisällä voivat heikentää luokittelumenetelmien tehokkuutta (Chen ym. 2004, Morgan ym. 2010).

Lasketuista kuvapiirteistä erottuivat selvästi sinisen kanavan arvot sekä siitä lasketut indeksit ja piirteet. Etenkin kuolleille puille niistä muodostui tärkeimpiä selittäviä muuttujia. Sylvain ym. (2019) saivat mielenkiintoisia tuloksia, jotka vahvistivat osaltaan sinisen kanavan hyötyä kuolleiden puiden kartoituksessa. He vertasivat omassa tutkimuksessaan kolmea eri kanavakokoonpanoa (RGBI Red-Green-Blue-Infrared, RGB Red-Green-Blue sekä IRG Infrared-Red-Green) luokittelun tarkkuudessa ja tuloksista huomattiin huonoimmin pärjänneen vaihtoehdon, joka ei sisältänyt olleenkaan sinistä kanavaa (IRG). Huonoimmassa vaihtoehdossa lähi-infrapun kanava korvasi sinisen kanavan, ja vaikka lähi-infrapunasäteily vähenee merkittävästi kuolleilla puilla (Haara ja Nevalainen 2002, Heikkilä ym. 2002) ei se kyennyt korvaamaan sinistä kanavaa.

LR näytti lisäksi käytettyjen piirteiden merkitsevyyden luokittain ja tuloksista nähtiin sinisen kanavan piirteiden linkittyvän kuolleisiin puihin ollen niiden tärkeimpiä selittäviä piirteitä. Hieman yllättäen aiemmista tutkimuksista tuttu ja kuolleiden puiden luokittelussa hyväksi todettu punaisen ja vihreän kanavan indeksi *RGI* (Red-Green index, Coops ym. 2006, Meddens ym. 2011, Hart ja Veblen 2015) ja siitä lasketut tilastolliset muuttujat eivät erottuneet joukosta. Vähemmän yllättäen sairaiden puiden tärkeimmät piirteet liittyivät punaiseen kanavaan. Siltikään *RGI* ja siitä lasketut muuttujat eivät osoittautuneet tärkeimmiksi, vaan punaisen ja vihreän kanavan erotuksesta lasketut tilastolliset muuttujat, etenkin persentiilit, kuten erotuksen 90. tai 95. persentiili (R_G_p90 , R_G_p95) olivat parhaita. Terveillä puilla vaihtelua piirteissä oli eniten, mutta vihreästä kanavasta lasketut tilastolliset muuttujat, kuten maksimiarvo (G_max) ja keskihajonta (G_std) toistuivat joka vuosi. Lisäksi vihreän ja sinisen kanavan (*GBI*) sekä vihreän ja punaisen kanavan (*GRI*) indekseistä lasketut muuttujat erottuivat joukosta.

Ortoilmakuvilta määritetyn puun sijainti saattoi poiketa puun todellisesta sijainnista. Keskusprojektion aiheuttama projektiovirhe saa kohteet näkymään ortoilmakuvissa vinossa eikä välttämättä kohtisuoraan ylöspäin. Puiden sijainti määritettiin massan keskikohtaan eikä latvan huippuun. Näin ollen kuvalta määritetty puun sijainti ja todellinen sijainti saattoivat poiketa toisistaan. Tämä taas mahdollistaa puuaineiston ja laserilta saadun latvuskartan yhdistämisessä tapahtuvia virheitä. Sijainnin ollessa riittävän vinossa puu voi saada viereisen puun latvussegmentin eikä sille oikeasti kuuluvaa latvusta. Virheellisen latvussegmentin alueella kuvassa näkyvä kohde voi olla vääränlainen luokiteltavaan puuhun verratessa. Myös huonolaatuiset latvussegmentit aiheuttivat ongelmia (Kuva 15).

Näin aineistoon syntyi ”virheellisiä” puita. Lisäksi aineistojen ajallinen eroavaisuus osaltaan vahvisti tätä ongelmaa. Näitä puita yritettiin tarkastella tuloksista vertaamalla kaikkien menetelmien ennusteita referenssipuuhun. Jos kaikki menetelmät ennustivat saman luokan, mutta se ei ollut oikea oletettiin kyseessä olevan mahdollisesti virheellinen puu. Keskimäärin kaikkina vuosina näiden virheellisten puiden osuus kaikista luokitelluista puista oli noin 2 %, mutta 2015 määrä oli 5 %. Tästä voitiin osittain tulkita 2015 vuoden huonon tuloksen johtuneen mahdollisesti kuvan tavallista suuremmista projektiovirheistä ja niistä syntyneistä virheellisistä puista aineistossa.



Kuva 15. Vasemmalla kuolleelle puulle muodostettu liian iso latvussegmentti sisältäen useita eläviä puita ortokuvasta katsottuna. Oikealla sama kohta CHM:n näkökulmasta.

4.3 Luokittelumenetelmien erojen tarkastelu

Kaikkina vuosina molemmilla mittareilla selvästi paras menetelmä oli LR. LR:n tulokset olivat johdonmukaisia ja hyvällä tasolla vuodesta riippumatta. Lisäksi sen tulokset erottuivat muista menetelmistä, vaikka esimerkiksi vuoden 2017 tulokset ovat hyvin tasaisia kolmen muun menetelmän kesken. Kyseisen menetelmän on todettu ennenkin onnistuvan hyvin kuolleiden puiden luokittelussa. Krzystek ym. (2020) käyttivät omassa tutkimuksessaan samaista logistista regressiota kuolleiden puiden luokitteluun, sillä heidän kokeilujensa mukaan se antoi parempia tuloksia verrattuna esimerkiksi Random Forestiin.

LDA:n luokittelutarkkuudet olivat neljänä vuotena huonoimmat ja etenkin vuoden 2005 tulokset poikkesivat merkittävästi muista saman vuoden tuloksista. Kuolleiden puiden oikeinluokittelutarkkuus (84 %) oli poikkeuksellisen huono myös muiden vuosien luokitteluihin verrattessa. Kokonaisuutena LDA:n tulokset olivat kaksijakoisia, sillä heikkojen vuosien lisäksi se oli kolmena vuotena toiseksi paras luokittelutarkkuuksia tarkastellessa. Kappa-arvoilla mitattuna LDA oli toiseksi paras neljä kertaa, joten tuloksissa on enemmän vaihtelua, kuin muilla luokittelumenetelmillä. LDA on toimintaperiaatteiltaan yksinkertainen luokittelumenetelmä, joten osittain tulokset voivat johtua suuresta piirteiden määrästä sekä vaikeudesta löytää kolmen luokan välisiä eroja.

Random Forest -pohjaisten menetelmien tuloksissa oli vain pieniä eroja keskenään. Sekä luokittelutarkkuudet että kappa-arvot olivat melkein jokaisena vuotena lähes samoja. Menetelmien keskinäisessä paremmuusjärjestyksessä oli hieman vaihtelua. Lähimmän naapurin menetelmä (RF+kNN) oli luokittelutarkkuuksissa sekä paras että myös heikoin, kun taas RF oli lähes jokaisena vuonna kolmas. Kokonaisuutena tuloksiin ei näyttänyt vaikuttavan kumpaa Random Forest -menetelmää käytti, sillä kummallakaan ei ole tulosten perusteella suuria etuja tai haittoja verrattuna toiseen.

Taulukosta 9 nähdään oikeinluokitteluprosentit sekä vuosien että luokittelumenetelmien välillä. Lisäksi tuloksista nähdään terveiden puiden luokittelutarkkuuksien olleen lähes kaikissa tapauksissa kuolleiden puiden vastaavia parempia. Lasketuista konfuusiomatriiseista voitiin myös tarkastella mihin luokkiin väärin luokitellut puut useimmiten päättyivät. Terveet puut luokiteltiin väärin valtaosassa tapauksista kuolleiksi. Sairaant puut taas luokiteltiin väärin lähes poikkeuksetta kuolleiksi puiksi. Kuolleiden puiden luokittelussa

esiintyi enemmän vaihtelua ja vuodesta riippuen suurin osa luokiteltiin väärin terveeksi tai väärin luokitellut jakautuivat melko tasaisesti sekä terveisiin että sairaisiin puihin.

Lisäksi taulukosta 9 voidaan tarkastella eri menetelmien oikeinluokittelutarkkuuksia luokittain. Muutamia poikkeuksia lukuun ottamatta terveiden ja kuolleiden puiden tarkkuudet olivat tasaisia eri menetelmien välillä. Suurimmat erot syntyivät sairaiden puiden luokittelussa. Neljänä vuotena LDA:n sairaiden puiden oikeinluokittelutarkkuus oli paras, mutta vuonna 2019 se oli huonoin. Vuoden 2019 aineistossa oli kappalemäärällisesti eniten sairaita puita (45), mikä auttoi muita luokittelumenetelmiä. Näyttikin siltä, että LDA ei kyennyt riittävän tehokkaasti erottelemaan kolmea eri luokkaa toisistaan, kun yksikään luokista ei ollut selvästi pienin.

Tuloksista nähdään myös luokitteluaineiston luokkien keskinäisten suhteiden vaikutuksia. Vuoden 2019 aineisto oli sairaiden puiden suurimman lukumäärän lisäksi tasapainoisin luokkien välillä. Sen seurauksen kaikkien luokkien tarkkuudet olivat hyvällä tasolla. Vuoden 2011 aineistossa sairaita puita oli toiseksi eniten (42) ja tuloksissa näkyy positiivisesti aineiston luokkien välinen tasapaino. Luokitteluaineiston suuri vinous luokkaa tai luokkia kohtaan vaikuttikin suoraan luokittelutulokseen ennusteiden painottuessa määrällisesti suurimpiin luokkiin.

5. JOHTOPÄÄTÖKSET

Tutkielman tavoitteena oli selvittää, miten puiden kuolleisuus on vaihdellut tutkimusalueella määrällisesti ja ajallisesti. Lisäksi selvitettiin millä tarkkuudella kuolleita puita voidaan tunnistaa kaukokartoitusaineistoilta. Tuloksista nähtiin selviä muutoksia puiden kuolleisuudessa tutkimusjakson aikana. Ilmakuva-aikasarjaan perustuva visuaalinen kuolleiden puiden tunnistaminen osoittautui toimivaksi menetelmäksi. Kyseisellä tavalla voidaan tuottaa kohtuullisella työllä ja suurella tarkkuudella hyödyllisiä aikasarjoja kuvaamaan kuolleiden puiden määriä alueellisesti. Manuaalinen kuvatulkinna mahdollistaa myös uusien kuolleiden puiden kartoittamisen, mikä on erittäin hyödyllinen tieto seurattaessa puiden kuolleisuutta jatkuvan tuhon alueilla.

Monivuotisella seurannalla voidaan havaita muutoksia puutasolla ja ennustaa muuttuvien ilmasto-olosuhteiden vaikutuksia paremmin. Ilmastonmuutoksen seurauksena vuosien keskilämpötilojen odotetaan nousevan ja keskilämpötilojen sekä puiden kuolleisuuden välillä havaittiin positiivinen riippuvuus. On kuitenkin vaikea arvioida ilmastomuuttujien todellisia vaikutuksia kuolleisuuteen, mutta yhdistämällä tarkat säähavainnot muutoksiin yksittäisten puiden terveydentilassa voidaan tulevaisuudessa ymmärtää paremmin ilmastomuutoksen vaikutuksia.

Tutkielman toisen puoliskon tuloksista nähtiin saavutettujen luokittelutarkkuuksien olleen vähintäänkin hyviä. Tulokset olivat samalla tasolla aiempien samankaltaisten tutkimusten kärkipään kanssa. Lisäksi voidaan huomioda tässä tutkielmassa käytetty aineisto. Kaikissa vastaavissa tutkimuksissa luokitteluissa käytetty aineisto on ollut laajempaa sisältäen useampia optisen kuva-aineiston kanavia tai mahdollisesti hyödyntänyt myös laserpiirteitä. Näin ollen saavutettuja luokittelutarkkuuksia pelkillä RGB-kuvilla voidaan pitää erittäin hyvinä. Tulokset saatiin laajahkolla aineistolla kaiken kaikkiaan seitsemänä vuotena, mikä lisää niiden luotettavuutta ja ne ovat yleistettävissä laajemmille alueille.

Puoliautomaattisista koneoppimismenetelmistä käytettiin neljää eri algoritmia. Kaikki algoritmit kykenivät vähintäänkin kohtuulliseen automaattiseen piirrevalintaan isosta aineistosta. Aineiston tulisi myös olla mahdollisimman tasainen luokkien välillä, etenkin sairaiden puiden luokittelu kärsi osaltaan vuosina, jolloin aineisto oli epätasapainoinen. Algoritmit oppivat aineiston sisällä luokkien väliset suhteet ja tuloksista nähtiin vinon jakauman vaikutus tuloksiin.

Kehittyneet koneoppimismenetelmät kykenevät jo nyt erittäin hyviin luokittelutarkkuuksiin. Suuremman mittakaavan käyttöönoton esteeksi muodostuvatkin saatavilla olevat kaukokartoitusaineistot sekä niiden yhteensovittaminen. Ortoilmakuvien tarkkuus on jo riittävällä tasolla, eikä spatiaalisen resoluution tarkkuudella ollut suurta merkitystä luokittelun onnistumiseen. Enemmän vaikuttaakin ortokuvien laatu, kuvaus tulisi suorittaa mahdollisimman tiheillä kuvanottoväleillä, jotta lopputuloksessa olisi mahdollisimman vähän projektiovirheitä. Dronet mahdollistavat laadukkaat ortokuvat, tosin vielä tällä hetkellä vain pienialaisesti. Latvuskartan muodostaminen laserkeilausaineistosta, ja yksittäisten puiden luokittelu latvusrajauksen perusteella on hyvin yleistettävissä isoillekin alueille. Yksinpuintulkinta ei kuitenkaan ole virheetöntä, ja vaihtelevissa metsiköissä

sekä puiden tunnistaminen että myös latvuksen rajausta vaatii vielä kehittämistä. Etenkin rajattujen latvusten tulisi todellakin sisältää vain yhden puun latva eikä mahdollisesti useamman puun latvoja tai osia niistä.

Toteutetun tutkielman perusteella voidaan perustellusti todeta, että ilmakehän aikasarjoista muodostetut puiden kuolleisuuden aikasarjat kuvaavat kuolleiden puiden määriä ja sijainteja hyvin. Lisäksi kuolleiden puiden luokittelun pelkillä RGB-kuvilla voidaan sanoa olevan riittävän tarkkaa laajamittaiseen käyttöön.

6. LÄHTEET

- Annala, E. 2001. Kaarnakuoriaistuhot vältettävissä hyvällä metsänhoidolla. *Metsätieteen aikakauskirja* 2/2001: 265–269. <https://doi.org/10.14214/ma.6501>.
- Ayres, M.P. & Lombardero, M.J. 2000. Assessing the consequences of global change for forest disturbance from herbivores and pathogens. *Science of the Total Environment* 262(3): 263–286.
- Boehnel, H.J. 1976. Spectral field measurements of agricultural and forest crops. In: *Proceedings of the XIV IUFRO Worlds Congress*. Oslo, 6 June-7 July 1976, s. 110–119.
- Breiman, L. 2001. Random Forests. *Machine learning* 45: 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Bright, B.C., Hudak, A.T., Egan, J.M., Jorgensen, C.L., Rex, F.E., Hicke, J.A. & Meddens, A.J.H. 2020. Using Satellite Imagery to Evaluate Bark Beetle-Caused Tree Mortality Reported in Aerial Surveys in a Mixed Conifer Forest in Northern Idaho, USA. *Forests* 2020 11(5): 529. <https://doi.org/10.3390/f11050529>.
- Brockhaus, J., Campbell, M., Bruck, R. & Khorram, S. 1989. Analysis of forest decline in the Southern Appalachians with TM data. In: *Proceedings of the ASPRS/ACSM 1989 Annual Convention Agenda for the 90's*. Baltimore 2-7 April 1989. American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, s. 227–236.
- Chen, X.X., Vierling, L., Rowell, E. & DeFelice, T. 2004. Using lidar and effective LAI data to evaluate IKONOS and Landsat 7 ETM+ vegetation cover estimates in a ponderosa pine forest. *Remote Sensing of Environment* 91(1): 14–26. <https://doi.org/10.1016/j.rse.2003.11.003>.

- Christiansen, E. & Bakke, A. 1988. The spruce bark beetle of Eurasia. In: Berryman, A. A. (ed.). Dynamics of forest insect populations: Patterns, Causes, Implications. Plenum Press, New York. s. 480–503.
- Citynature 2020. Keskuspuisto. [Verkkodokumentti]. Saatavissa: <https://citynature.eu/fi/kohde/keskuspuisto/?section=nature>. [Viitattu 6.1.2021].
- Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20(1): 37–46.
<https://doi.org/10.1177/001316446002000104>.
- Colwell, J.E. 1970. Applications of remote sensing in agriculture and forestry. Remote Sensing with Special Reference to Agriculture and Forestry. National Academy of Sciences, Washington DC. s. 164–223.
- Coops, N.C, Johnson, M., Wulder, M.A. & White, J.C. 2006. Assessment of QuickBird high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation. Remote Sensing of Environment 103(1): 67–80.
<https://doi.org/10.1016/j.rse.2006.03.012>.
- Crookston, N.L. & Finley, A.O. 2008. yaImpute: An R Package for kNN Imputation. Journal of Statistical Software 23(10): 1–16. <https://doi.org/10.18637/jss.v023.i10>.
- Forest.fi 2014. Helsinki's Central Park is really central forest. [Verkkodokumentti]. Saatavissa: <https://forest.fi/article/helsinkis-central-park-is-really-central-forest/>. [Viitattu 3.4.2020].
- Franklin, S.E., Wulder, M.A., Skakun, R.S. & Carroll, A.L. 2003. Mountain pine beetle red-attack forest damage classification using stratified Landsat TM data in British Columbia, Canada. Photogrammetric Engineering and Remote Sensing 69(3): 283–288. <https://doi.org/10.14358/PERS.69.3.283>.
- Freeman, M.P., Stow, D.A. & Roberts, D.A. 2016. Object-based Image Mapping of Conifer Tree Mortality in San Diego County based on Multitemporal Aerial Ortho-

imagery. *Photogrammetric Engineering & Remote Sensing* 82(7): 571–580.
<https://doi.org/10.14358/PERS.82.7.571>.

Friedman, J., Hastie, T., Tibshirani, R., Narasimhan, B., Tay, K., Simon, N. & Qian, J. 2021. Lasso and Elastic-Net Regularized Generalized Linear Models. [Verkkodokumentti]. Saatavissa:
<https://cran.r-project.org/web/packages/glmnet/glmnet.pdf>. [Viitattu 21.2.2021].

Gitelson, A.A., Kaufman, Y.J., Stark, R. & Rundquist, D. 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment* 80(1): 76–87.

Haara, A. & Nevalainen, S. 2002. Detection of dead or defoliated spruces using digital aerial data. *Forest Ecology and Management* 160(1–3): 97–107.

Haggrén, H. 2002. Ortokuvien tuottaminen. [Verkkodokumentti]. Saatavissa:
<https://foto.aalto.fi/opetus/220/luennot/7/7.html>. [Viitattu 18.2.2021].

Hart, S.J. & Veblen, T.T. 2015. Detection of spruce beetle-induced tree mortality using high- and medium-resolution remotely sensed imagery. *Remote Sensing of Environment* 168: 134–145. <https://doi.org/10.1016/j.rse.2015.06.015>.

Hast, A. 2010. Diskriminantianalyysi I. [Verkkodokumentti]. Saatavissa: http://sal-server.org.aalto.fi/vanhat_sivut/Opinnot/Mat-2.4142/S2010/esitykset/es19_aira.pdf. [Viitattu 14.1.2021].

Hastie, T., Qian, J. & Tay, K. 2021. An introduction to glmnet. [Verkkodokumentti]. Saatavissa: <https://glmnet.stanford.edu/articles/glmnet.html>. [Viitattu 14.1.2021].

Heikkilä, J., Nevalainen, S. & Tokola, T. 2002. Estimating defoliation in boreal coniferous forests by combining Landsat TM, aerial photographs and field data. *Forest Ecology and Management* 158(1–3): 9–23.

- Helsingin kaupunki 2020. Keskuspuisto metsä lähellä – luonto kaupungissa. [Verkkodokumentti]. Saatavissa: <https://www.hel.fi/hel2/keskuspuisto/fin/1esittely/>. [Viitattu 24.3.2020].
- Helsingin paikkatietohakemisto 2021. [Verkkodokumentti]. Saatavissa: <https://kartta.hel.fi/paikkatietohakemisto/pti/>. [Viitattu 6.1.2021].
- Holopainen, M., Hyypä, J. & Vastaranta M. 2013. Laserkeilaus metsävarojen hallinnassa. Helsingin yliopiston metsätieteiden laitoksen julkaisuja 5. 75 s. <http://hdl.handle.net/10138/42935>.
- Holopainen, M., Tokola, T., Vastaranta, M., Heikkilä, J., Huitu, H., Laamanen, R. & Alho, P. 2015. Geoinformatiikka luonnonvarojen hallinnassa. Helsingin yliopiston metsätieteiden laitoksen julkaisuja 7. 152 s. <http://hdl.handle.net/10138/166765>.
- Howard, J.A. 1991. Remote Sensing of Forest Resources. Theory and Application. Chapman & Hall, London. 420 s.
- Hudak, A., Crookston, N., Evans, J., Hall, D. & Falkowski, M. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment 112(5): 2232–2245. <https://doi.org/10.1016/j.rse.2007.10.009>.
- Hyypä, J. & Inkinen, M. 1999. Detecting and estimating attributes for single trees using laser scanner. The Photogrammetric Journal of Finland 16(2): 27–42.
- Ilmastokatsaus 2019. Lämpenevä ilmasto suosii metsätuholaisia. [Verkkodokumentti]. Saatavissa: <http://www.ilmastokatsaus.fi/2019/06/20/lampeneva-ilmasto-suosii-metsatuholaisia/>. [Viitattu 18.2.2021].
- Ilmatieteen laitos 2006. Tiedote – Ennätysellisen kuiva hellekesä päättymässä. [Verkkodokumentti]. Saatavissa: <https://www.ilmatieteenlaitos.fi/tiedote/1157014828>. [Viitattu 20.3.2021].

- Ilmatieteen laitos 2020. Ilmastokatsaus – talvi oli ennätysellisen leuto. [Verkkodokumentti]. Saatavissa: <http://www.ilmastokatsaus.fi/2020/03/23/talvi-oli-ennatysellisen-leuto/>. [Viitattu 4.4.2020].
- Ilmatieteen laitos 2021. Lämpötila- ja sadetilastoja vuodesta 1961. [Verkkodokumentti]. Saatavissa: <https://www.ilmatieteenlaitos.fi/tilastoja-vuodesta-1961>. [Viitattu 20.3.2021].
- Junttila, S. 2019. Utilizing multispectral lidar in the detection of declined trees. *Dissertationes Forestales* 277. 58 s. <https://doi.org/10.14214/df.277>.
- Jönsson, A.M., Harding, S., Barring, L. & Ravn, H.P. 2007. Impact of climate change on the population dynamics of *Ips typographus* in southern Sweden. *Agricultural Forest Meteorology*. 146(1–2): 70–81. <https://doi.org/10.1016/j.agrformet.2007.05.006>.
- Kankaanhuhta, V., Pouttu, A., Neuvonen, S., Viiri, H. & Siitonen, J. 2014. MetINFO – metsien terveys, Kirjanpainaja (*Ips typographus*). [Verkkodokumentti]. Luonnonvarakeskus. Saatavissa: http://www.metla.fi/metinfo/metsienterveys/lajit_kansi/iptypo-n.htm. [Viitattu 4.4.2020].
- Kellner, J.R. & Hubbell, S.P. 2017. Adult mortality in a low-density tree population using high-resolution remote sensing. *Ecology* 98(6): 1700–1709. <https://doi.org/10.1002/ecy.1847>.
- Kinnunen, R., Lehtonen, I., Kas, J., Järvelä, R., Poutamo, H., Wenzlaff, C. & Latus, J. 2013. Impact of Climate Change on the Boreal Forest in Finland and Sweden. University of Helsinki. 34 s. Saatavissa: <https://www.researchgate.net/publication/271271260>.

- Krzystek, P., Serebryanyk, A., Schnörr, C., Červenka, J. & Heurich, M. 2020. Large-Scale Mapping of Tree Species and Dead Trees in Šumava National Park and Bavarian Forest National Park Using Lidar and Multispectral Imagery. *Remote Sens.* 2020, 12(4): 661. <https://doi.org/10.3390/rs12040661>.
- Kärvemö, S. & Schroeder, L.M. 2010. A comparison of outbreak dynamics of the spruce bark beetle in Sweden and the mountain bark beetle in Canada (Curulionidae: Scolytidae). *Entomologisk Tidskrift*. 131(3): 215–224.
- Kärvemö, S., Rogell, B. & Schroeder, M. 2014a. Dynamics of spruce bark beetle infestation spots: Importance of local population size and landscape characteristics after a storm disturbance. *Forest ecology and management*. 334: 232–240. <https://doi.org/10.1016/j.foreco.2014.09.011>.
- Kärvemö, S., Van Boeckel, T.P., Gilbert, M., Grégoire, J.C. & Schroeder, M. 2014b. Large-scale risk mapping of an eruptive bark beetle – Importance of forest susceptibility and beetle pressure. *Forest ecology and management*. 318: 158–166. <https://doi.org/10.1016/j.foreco.2014.01.025>.
- Landis, J.R. & Koch, G.G. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33(1): 159–174. <https://doi.org/10.2307/2529310>.
- Latifi, H., Fassnacht, F.E., Müller, J., Tharani, A., Dech, S. & Heurich, M. 2015. Forest inventories by LiDAR data: A comparison of single tree segmentation and metric-based methods for inventories of a heterogeneous temperate forest. *International Journal of Applied Earth Observation and Geoinformation* 42: 162–174. <https://doi.org/10.1016/j.jag.2015.06.008>.
- Long, J.A. & Lawrence, R.L. 2016. Mapping Percent Tree Mortality Due to Mountain Pine Beetle Damage. *Forest Science* 62(4): 392–402. <https://doi.org/10.5849/forsci.15-046>.
- Lundquist, J.E., Camp, A.E., Tyrell, M.L., Seybold, S.J., Cannon, P. & Lodge, D. 2011. Earth, wind, and fire: Abiotic factors and the impacts of global environmental

change on forest health. In: Castello, J.D. and Teale, S.A. (eds). *Forest Health: An Integrated Perspective*. Cambridge University Press, Cambridge, UK. s. 195–243.

Luonnonvarakeskus 2021a. Metsät ja ilmastonmuutos. [Verkkodokumentti]. Saatavissa: <https://www.luke.fi/tietoa-luonnonvaroista/metsa/metsat-ja-ilmastonmuutos/>. [Viitattu 20.3.2021].

Luonnonvarakeskus 2021b. Metsäinfo – Tuhonaiheuttajat. [Verkkodokumentti]. Saatavissa: <https://metsainfo.luke.fi/fi/cms/opas/tuhonaiheuttajaluettelo>. [Viitattu 18.2.2021].

Marini, L. 2012. Climate affects severity and altitudinal distribution of outbreaks in an eruptive bark beetle. *Climate Change* 115(2): 327–341.
<https://doi.org/10.1007/s10584-012-0463-z>.

McHugh, M.L. 2012. Interrater reliability: the kappa statistic. *Biochemia Medica* 22(3): 276–282. <https://doi.org/10.11613/BM.2012.031>.

Meddens, A.J.H., Hicke, J.A. & Vierling, L.A. 2011. Evaluating the potential of multi-spectral imagery to map multiple stages of tree mortality. *Remote Sensing of Environment* 115(7): 1632–1642. <https://doi.org/10.1016/j.rse.2011.02.018>.

Meddens, A.J.H., Hicke, J.A., Vierling, L.A. & Hudak, A.T. 2013. Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date Landsat imagery. *Remote Sensing of Environment* 132: 49–58.
<https://doi.org/10.1016/j.rse.2013.01.002>.

Meyer, F. & Beucher, S. 1990. Morphological segmentation. *Journal of Visual Communication and Image Representation* 1(1): 21–46.

Morgan, J.L., Gergel, S.E. & Coops, N.C. 2010. Aerial photography: A rapidly evolving tool for ecological management. *Bioscience* 60(1): 47–59.
<https://doi.org/10.1525/bio.2010.60.1.9>.

- My Helsinki 2020. Keskuspuisto – Helsinkiä halkova keskusmetsä. [Verkkodokumentti].
Saattavissa: <https://www.myhelsinki.fi/fi/n%C3%A4e-ja-koe/luonto/keskuspuisto-helsinki%C3%A4-halkova-keskusmets%C3%A4>. [Viitattu 31.3.2020].
- Naesset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* 52(2): 49–56.
- Nuorteva, H. (toim). 2019. Metsätuhot vuonna 2018. Luonnonvara- ja biotaloudentutkimus 85/2019. Luonnonvarakeskus. Helsinki. 60s. <http://urn.fi/URN:ISBN:978-952-326-878-4>.
- Näsi, R., Oliveira, R., Viljanen, N., Suomalainen, J., Markelin, L., Hakala, T. & Honkavaara, E. 2020. Metsien terveydentilan määrittäminen droonipohjaisella kaukokartoituksella. Maanmittauslaitos, Paikkatietokeskus (FGI).
- Parviainen, J., Vapaavuori, E. & Mäkelä, A. 2010. Finland's Forests in Changing Climate. Working Papers of the Finnish Forest Research Institute 159: 1–50. Saatavissa: <http://www.metla.fi/julkaisut/workingpapers/2010/mwp159.htm>.
- Polewski, P. 2017. Reconstruction of Standing and Fallen Single Dead Trees in Forested Areas from LiDAR Data and Aerial Imagery. Ph.D. Thesis, Technische Universität München, München, Germany.
- Polewski, P., Yao, W., Heurich, M., Krzystek, P. & Stilla, U. 2015. Free Shape Context descriptors optimized with genetic algorithm for the detection of dead tree trunks in ALS point clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* II-3/W5: 41–48. <https://doi.org/10.5194/isprsannals-II-3-W5-41-2015>.
- Popescu, S.C. & Wynne, R.H. 2004. Seeing the trees in the forest. *Photogrammetric Engineering and Remote Sensing* 70(5): 589–604.
<https://doi.org/10.14358/PERS.70.5.589>.

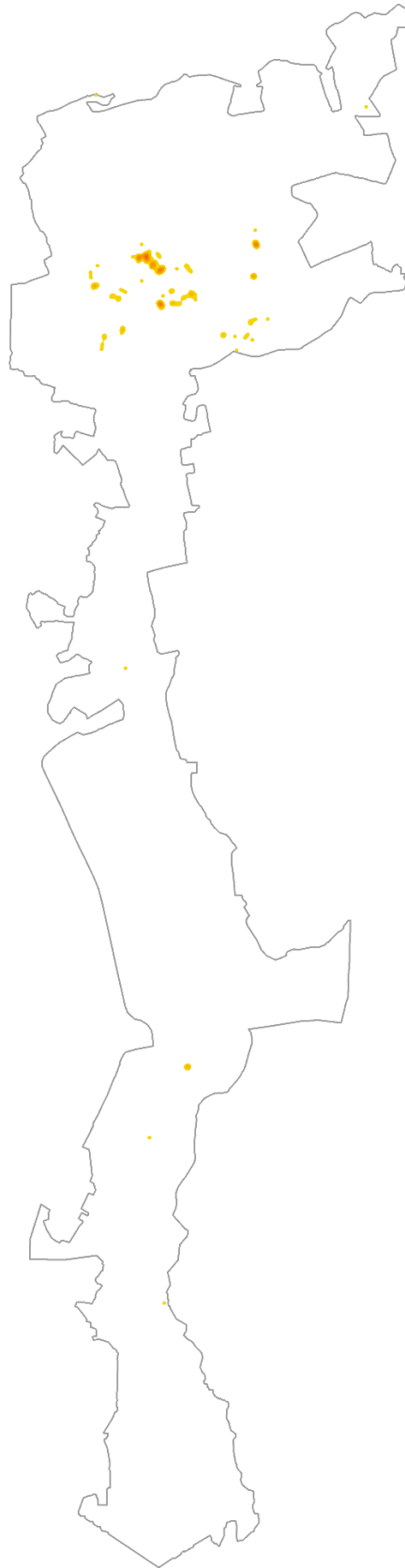
- Popescu, S.C., Wynne, R.H. & Nelson, R.F. 2002. Estimating plot-level tree heights with lidar: Local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture* 37(1–3): 71–95.
- Pouttu, A. & Annala, E. 2010. Kirjanpainajalla kaksi sukupolvea kesällä 2010. *Metsätieteen aikakauskirja* 4/2010: 521–523. <https://doi.org/10.14214/ma.6951>.
- R Core Team 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. [Verkkodokumentti]. Saatavissa: <https://www.r-project.org/>. [Viitattu 17.2.2021].
- Retkipaikka 2016. Helsingin keskuspuistossa on metsän tuntua. [Verkkodokumentti]. Saatavissa: <https://retkipaikka.fi/helsingin-keskuspuistossa-on-metsan-tuntua/>. [Viitattu 6.1.2021].
- Riley, J.R. 1989. Remote sensing in entomology. *Annual Review of Entomology* 34: 247–271. <https://doi.org/10.1146/annurev.en.34.010189.001335>.
- Roussel, J-R., Auty, D., De Boissieu, F., Sanchez Meador, A., Bourdon, J-F. & Gatzolis, D. 2021. Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. [Verkkodokumentti]. Saatavissa: <https://cran.r-project.org/web/packages/lidR/lidR.pdf>. [Viitattu 17.2.2021].
- Schroeder, M. & Lindelöw, Å. 2002. Attacks on living spruce trees by the bark beetle *Ips typographus* (Col. Scolytidae) following a storm-felling: a comparison between stands with and without removal of wind-felled trees. *Agricultural and Forest Entomology* 4(1): 47–56. <https://doi.org/10.1046/j.1461-9563.2002.00122.x>.
- Sirén, T. 2008. Yksimielisyyden mittaamisen tilastolliset menetelmät. Tampereen yliopisto matematiikan ja tilastotieteen laitos. Pro gradu -tutkielma. <http://urn.fi/urn:nbn:fi:uta-1-19384>.

- Skakun, R.S., Wulder, M.A. & Franklin, S.E. 2003. Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sensing of Environment* 86(4): 433–443.
- Sylvain, J-D., Drolet, G. & Brown, N. 2019. Mapping dead forest cover using a deep convolutional neural network and digital aerial photography. *ISPRS Journal of Photogrammetry and Remote Sensing* 156: 14–26.
<https://doi.org/10.1016/j.isprsjprs.2019.07.010>.
- Van Gunst, K.J., Weisberg, P.J., Yang, J. & Fan, Y. 2016. Do denser forests have greater risk of tree mortality: A remote sensing analysis of density-dependent forest mortality. *Forest Ecology and Management* 359: 19–32.
<https://doi.org/10.1016/j.foreco.2015.09.032>.
- Vastaranta, M., Kantola, T., Lyytikäinen-Saarenmaa, P., Holopainen, M., Kankare, V., Wulder, M.A., Hyyppä, J. & Hyyppä, H. 2013. Area-Based Mapping of Defoliation of Scots Pine Stands Using Airborne Scanning LiDAR. *Remote Sens.* 2013, 5(3): 1220–1234. <https://doi.org/10.3390/rs5031220>.
- Vogelman, J. & Rock, B. 1986. Assessing forest decline in coniferous forests of Vermont using NS-001 Thematic Mapper simulator data. *International Journal of Remote sensing* 7(10): 1303–1321. <https://doi.org/10.1080/01431168608948932>.
- Vosselman, G. & Maas, H-G. (eds.) 2010. *Airborne and terrestrial laser scanning*. CRC Press. 318 s.
- Wang, H., Zhao, Y., Pu, R. & Zhang, Z. 2016. Mapping Robinia Pseudoacacia forest health conditions by using combined spectral, spatial and textural information extracted from Ikonos imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B8: 1425–1429.
<https://doi.org/10.5194/isprs-archives-XLI-B8-1425-2016>.

- Wermelinger, B. 2004. Ecology and management of the spruce bark beetle *Ips typographus* – a review of recent research. *Forest Ecology and Management* 202(1–3): 67–82. <https://doi.org/10.1016/j.foreco.2004.07.018>.
- Wichmann, L. & Ravn, H.P. 2001. The spread of *Ips typographus* (L.) (Coleoptera, Scolytidae) attacks following heavy windthrow in Denmark, analysed using GIS. *Forest Ecology and Management* 148(1–3): 31–39.
- Wulder, M.A., Dymond, C.C., White, J.C., Leckie, D.G. & Carroll, A.L. 2006. Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. *Forest Ecology and Management* 221(1–3): 27–41. <https://doi.org/10.1016/j.foreco.2005.09.021>.
- Yao, W., Krzystek, P. & Heurich, M. 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment* 123: 368–380. <https://doi.org/10.1016/j.rse.2012.03.027>.

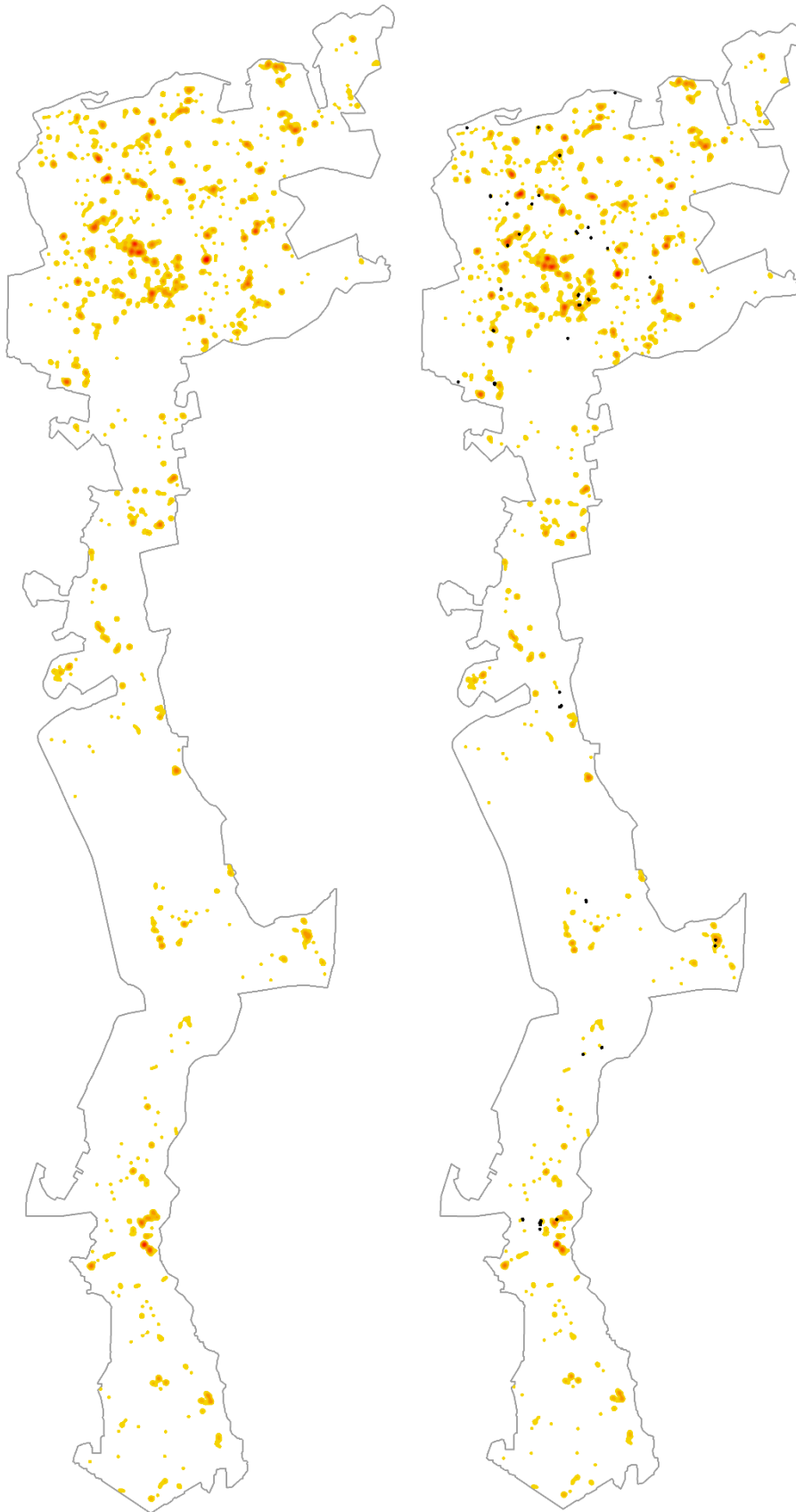
LIITTEET

2005



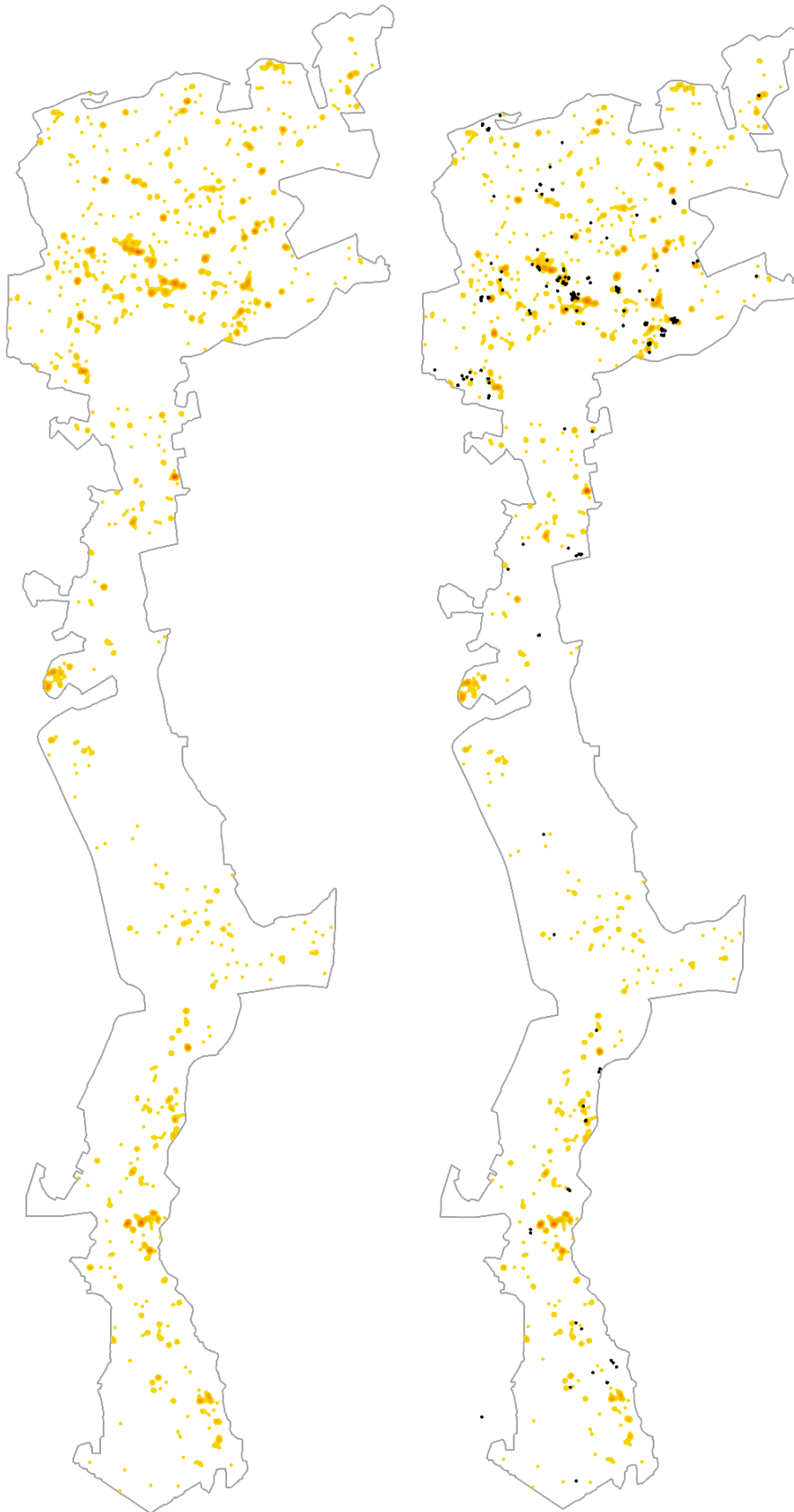
Kuva 1. Kuolleiden puiden sijainneista luotu lämpökarttakuva vuodelta 2005.

2009



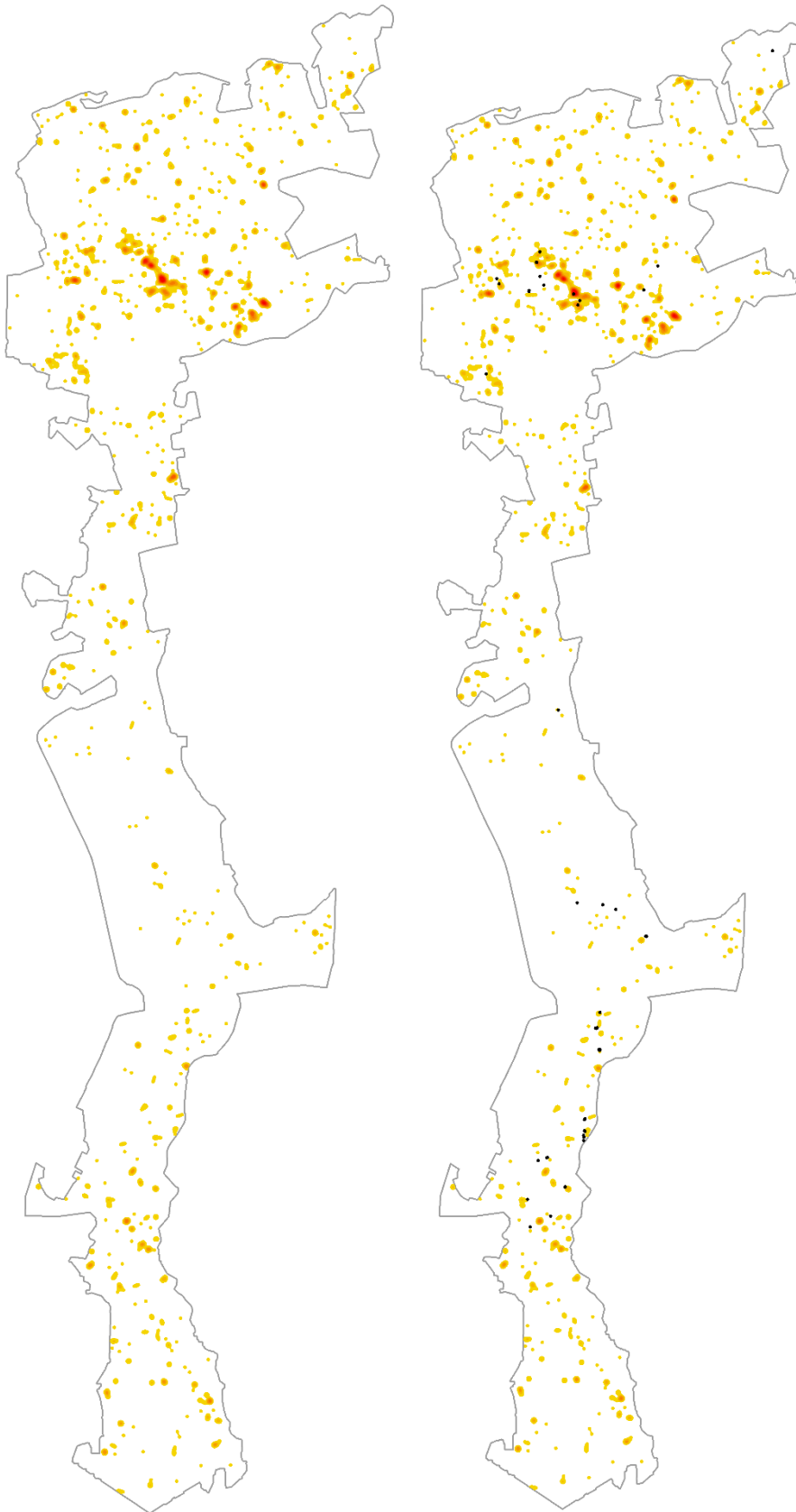
Kuva 2. Vasemmalla kuolleiden puiden sijainneista luotu lämpökarttakuva vuodelta 2009. Oikealla kuvaan on lisätty mustilla pisteillä sairaiden puiden sijainnit.

2011



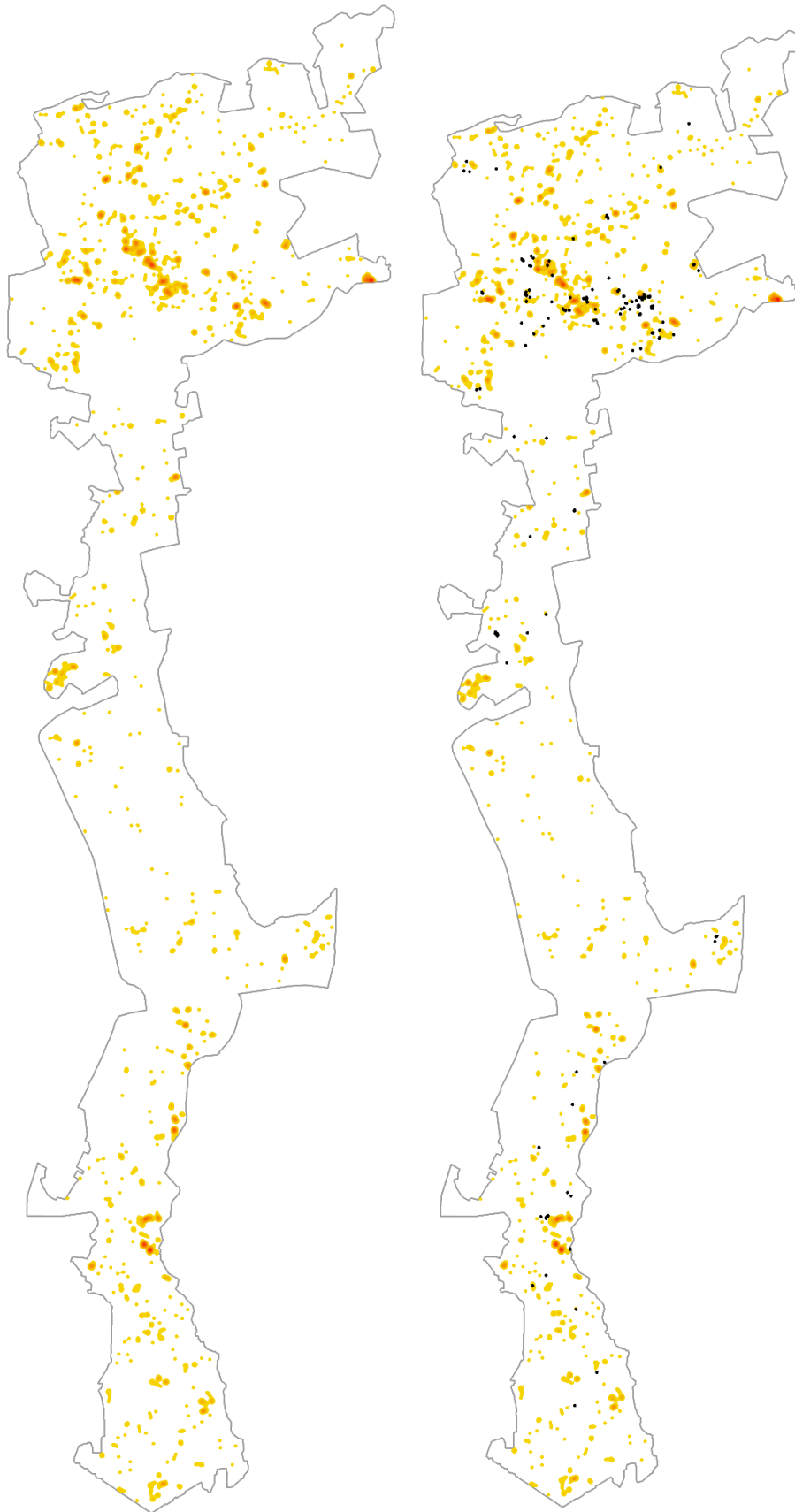
Kuva 3. Vasemmalla kuolleiden puiden sijainneista luotu lämpökarttakuva vuodelta 2011. Oikealla kuvaan on lisätty mustilla pisteillä sairaiden puiden sijainnit.

2013



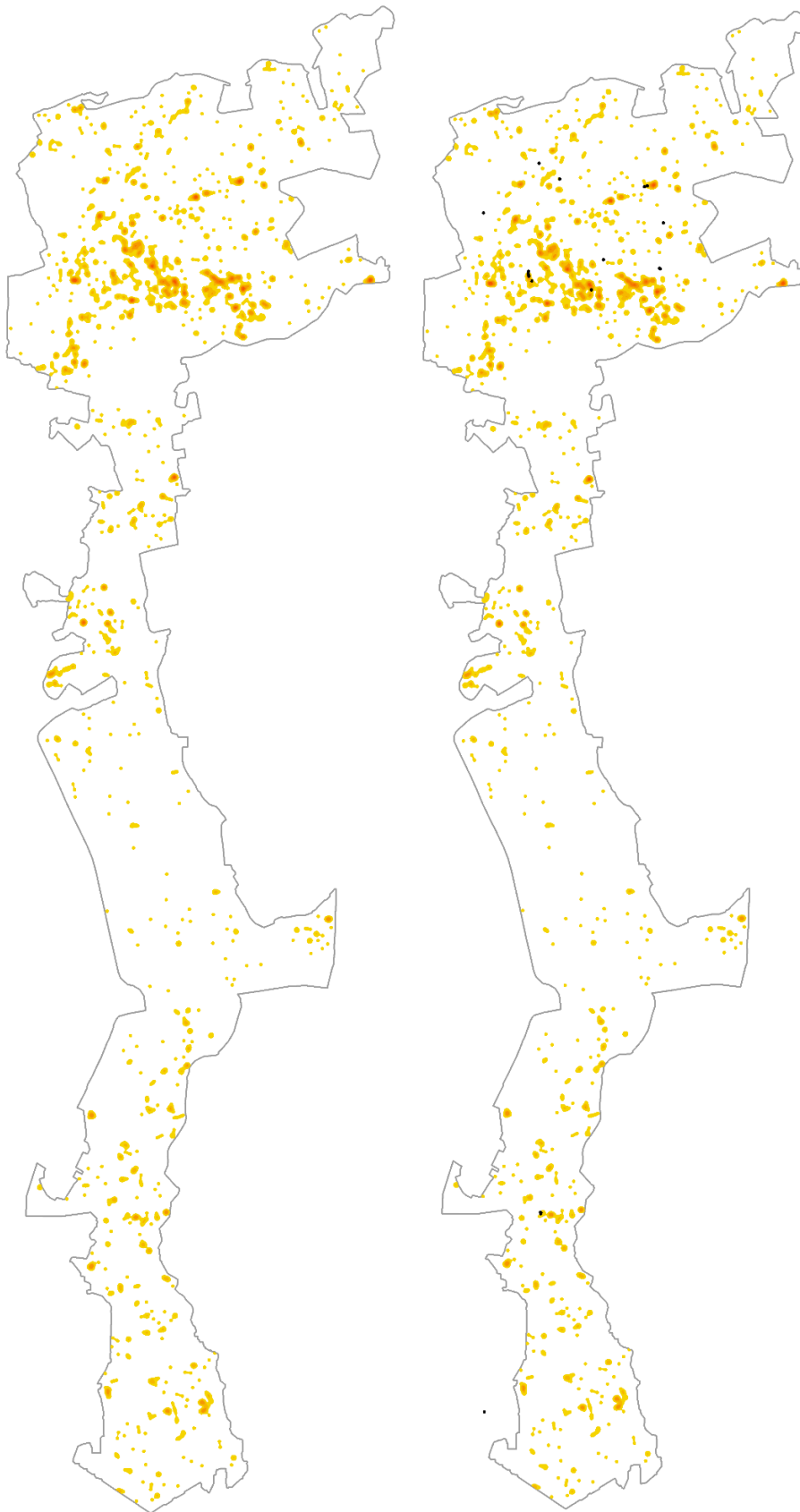
Kuva 4. Vasemmalla kuolleiden puiden sijainneista luotu lämpökarttakuva vuodelta 2013. Oikealla kuvaan on lisätty mustilla pisteillä sairaiden puiden sijainnit.

2015



Kuva 5. Vasemmalla kuolleiden puiden sijainneista luotu lämpökarttakuva vuodelta 2015. Oikealla kuvaan on lisätty mustilla pisteillä sairaiden puiden sijainnit.

2017



Kuva 6. Vasemmalla kuolleiden puiden sijainneista luotu lämpökarttakuva vuodelta 2017. Oikealla kuvaan on lisätty mustilla pisteillä sairaiden puiden sijainnit.